Advanced Techniques for Future Multicarrier Systems

Dipartimento di Ingegneria dell'Informazione: Elettronica, Informatica e Telecomunicazioni

Università di Pisa

A thesis submitted for the degree of

Dottore di Ricerca in Ingegneria dell'Informazione

January 2011

Author:

Ing. Giulio Dainelli

Advisors:

Prof. Aldo Nunzio D'Andrea

Ing. Marco Moretti

Prof. Michele Morelli

Acknowledgements

First of all I would like to thank my advisor, Marco Moretti, for his support. His help was crucial for my professional and personal growth. I have gained from his extraordinary motivation, great intuition and technical insight. Thanks to my other advisors, Professors Aldo D'Andrea and Michele Morelli, who have always been available when I needed them. I would also like to thank Professor Ruggero Reggiannini, with whom I had the fortune to work and collaborate. A special thanks goes to Alfredo Todini, for his invaluable advice and his constant availability.

Many thanks to my friends and colleagues of the Department of Information Engineering at the University of Pisa for their countless discussions and for a pleasant working atmosphere. I will always be indebted to my colleagues and friends of the Transmission Research Group for their helpful technical insights and valuable hints over these three years. Among these, Lorenzo Taponecco and Luca Sanguinetti have always been ready to listen to me and to have some fun together.

I would like to thank my parents for believing in me and for the sacrifices they made to get me here. If have written this thesis, a lot is due to them.

Finally, thanks to the most important person in my life, my wife Ilaria. I love you.

Pisa, February 2011 Giulio Dainelli

Abstract

Future multicarrier systems face the tough challenge of supporting high data-rate and high-quality services. The main limitation is the frequency-selective nature of the propagation channel that affects the received signal, thus degrading the system performance.

OFDM can be envisaged as one of the most promising modulation techniques for future communication systems. It exhibits robustness to ISI even in very dispersive environments and its main characteristic is to take advantage of channel diversity by performing dynamic resource allocation. In a multi-user OFDMA scenario, the challenge is to allocate, on the basis of the channel knowledge, different portions of the available frequency spectrum among the users in the systems.

Literature on resource allocation for OFDMA systems mainly focused on single-cell systems, where the objective is to assign subcarriers, power and data-rate for each user according to a predetermined criterion. The problem can be formulated with the goal of either maximizing the system sum-rate subject to a constraint on transmitted power or minimizing the overall power consumption under some predetermined constraints on rate per user. Only recently, literature focuses on resource allocation in multi-cell networks, where the goal is not only to take advantage of frequency and multi-user diversity, but also to mitigate MAI, which represents one of the most limiting factor for such problems.

We consider a multi-cell OFDMA system with frequency reuse distance equal to one. Allowing all cells to transmit on the whole bandwidth unveils large potential gains in terms of spectral efficiency in comparison with conventional cellular systems. Such a scenario, however, is often deemed unfeasible because of the strong MAI that negatively affects the system performance. In this dissertation we present a layered architecture that integrates a packet scheduler with an adaptive resource allocator, explicitly designed to take care of the multiple access interference. Each

cell performs its resource management in a distributed way without any central controller. Iterative resource allocation assigns radio channels to the users so as to minimize the interference. Packet scheduling guarantees that all users get a fair share of resources regardless of their position in the cell. This scheduler-allocator architecture integrates both goals and is able to self adapt to any traffic and user configuration. An adaptive, distributed load control strategy can reduce the cell load so that the iterative procedure *always* converges to a stable allocation, regardless of the interference. Numerical results show that the proposed architecture guarantees both high spectral efficiency and throughput fairness among flows.

In the second part of this dissertation we deal with FBMC communication systems. FBMC modulation is a valid alternative to conventional OFDM signaling as it presents a set of appealing characteristics, such as robustness to narrowband interferers, more flexibility to allocate groups of subchannels to different users/services, and frequency-domain equalization without any cyclic extension. However, like any other multicarrier modulations, FBMC is strongly affected by residual CFOs that have to be accurately estimated.

Unlike previously proposed algorithms, whereby frequency is recovered either relying on known pilot symbols multiplexed with the data stream or exploiting specific properties of the multicarrier signal structure in following a blind approach, we present and discuss an algorithm based on the ML principle, which takes advantage both of pilot symbols and also indirectly of data symbols through knowledge and exploitation of their specific modulation format. The algorithm requires the availability of the statistical properties of channel fading up to second-order moments. It is shown that the above approach allows to improve on both frequency acquisition range and estimation accuracy of previously published schemes.

Contents

Notation

$(\cdot)^T$	denotes transpose operation
$(\cdot)^H$	denotes Hermitian transposition
$[\cdot]_{k,\ell}$	denotes the (k,ℓ)th entry of the enclosed matrix
$(x)^+$	denotes the maximum between x and 0
$\|\cdot\|$	denotes the Euclidean norm
$\mathrm{tr}\{\cdot\}$	denotes the trace of a matrix
$\mathbf{I_N}$	denotes the identity matrix of order N
$\mathrm{card}\{\cdot\}$	denotes the cardinality of the enclosed set
$E\{\cdot\}$	denotes the expectation operator
$\mathbf{M}^{-1}$	denotes the inverse of a square matrix $\mathbf{M}$
$\mathrm{Re}\{\cdot\}$	denotes the real part of a complex-valued quantity
$\mathrm{Im}\{\cdot\}$	denotes the imaginary parts of a complex-valued quantity
$\|\cdot\|$	denotes the magnitude a complex-valued quantity

Acronyms

3GPP	Third Generation Partnership Project
AMC	Adaptive Modulation and Coding
AWGN	Additive White Gaussian Noise
BER	Bit Error Rate
BS	Base Station
CBFQ	Credit-Based Fair Queueing
CDMA	Code Division Multiple Access
CFO	Carrier Frequency Offset
CMLE	Combined Maximum Likelihood Estimator
CP	Cyclic Prefix
CRA	Centralized Multi-format Resource Allocator
CRLB	Cramer-Rao Lower Bound
CSI	Channel State Information
D/A	Digital-to-Analog
DA	Data-Aided
DD	Decision Directed
DFT	Discrete Fourier Transform
DLA	Dynamic Layered Architecture
DVB	Digital Video Broadcasting
EVD	Eigenvalue Decomposition
FBMC	Filter Bank Multicarrier
FFR	Fractional Frequency Reuse
FF	Fast-Fading
FIR	Finite Impulse Response

ICI	Inter Carrier Interference
IDFT	Inverse Discrete Fourier Transform
IPM	Interior Point Method
ISI	Inter Symbol Interference
JMLE	Joint Maximum Likelihood Estimator
LC- CMLE	Low Complexity Combined Maximum Likelihood Estimator
LC	Load Control
LIP	Linear Integer Programming
LLF	Log-Likelihood Function
LP	Linear Programming
LTE	Long Term Evolution
MAC	Media Access Control
MAI	Multiple Access Interference
MARA	Multi-Assign Resource Allocator
MA	Margin Adaptive
MLE	Maximum Likelihood Estimator
ML	Maximum Likelihood
MSEE	Mean Square Estimation Error
MS	Mobile Station
NP	Non-deterministic Polynomial time
NSM	Network Simplex Method
OFDMA	Orthogonal Frequency Division Multiple-Access
OFDM	Orthogonal Frequency Division Multiplexing
PA	Pilot-Aided
PSK	Phase Shift Keying
PS	Packet Scheduler
QAM	Quadrature Amplitude Modulation
RA	Rate Adaptive
RNC	Radio Network Controller
RRA	Radio Resource Allocator
SF	Slow-Fading
SINR	Signal to Interference plus Noise Ratio
SNR	Signal to Noise Ratio

SRRC	Square Root Raised Cosine
TETRA	Terrestrial Trunked Radio
TTI	Transmission Time Interval
TU	Typical Urban
UTRAN	Universal Terrestrial Radio Access Network
Wi-MAX	Worldwide Interoperability for Microwave Access
WLAN	Wireless Local Area Network

List of Figures

Chapter 1

Introduction

The rapid growth of wireless communications in the recent past has modified the way of communicating with each other. Digital cellular phones, as well as portable computer and fixed Internet technologies have contributed to the increase of wireless internet access. In many cases, such as in homes, offices and small urban areas, wireless local area networks have substituted wired networks. Many new applications, e.g. smart phones, wireless sensor networks, have been derived from research environments and produced in real systems. From the infrastructure point of view, the design of robust wireless communication network for each of these emerging applications represents the major challenge.

Future wireless communication systems will provide wideband access to large numbers of subscribers, while fulfilling at the same time strong requirements in terms of QoS. The challenge arises from the scarcity of frequency spectrum, the limitation on total transmit power and the nature of wireless channel. In wireless communications, the transmitted signals pass through a wireless propagation channel which is affected by frequency-selectivity and multipath fading, deriving from scattering, reflection, and diffraction of the radiated energy by objects in the environment or refraction in the medium. At the receiver side, the signal is the combination of different replicas of the original transmitted signal over each path. This leads to fluctuation of power of the received signal because of the constructive or destructive combination of the multipath components. Moreover, if the transmitter, receiver or surrounding objects are moving, the characteristic of the channel could change randomly. This leads to severe ISI both in time and frequency, thus degrading the data rate of communication. Dynamic resource allocation schemes that interact in both in the physical and the MAC layers are necessary to combat ISI.

1. INTRODUCTION

The idea is to allocate system resources (i.e. subcarriers, power and bit rate) to the users in the system according to the changing conditions of wireless propagation environment in order to achieve large gains in terms of the system capacity.

OFDM is one of the promising solutions to provide high performance from a physical layer point of view and it has been adopted in several standards, e.g. IEEE 802.11 (WLANs) [1], IEEE 802.16 (Wi-MAX) [2] and 3GPP LTE [3]. OFDM divides the whole bandwidth into N orthogonal narrowband subchannels, each with a bandwidth smaller than the coherence bandwidth of the channel. The stream at high data rate is split into N substreams with lower data rate, thus the N OFDM sysmbols are transmitted simultaneously on N orthogonal subcarriers. During each transmission block, each orthogonal subchannel can be approximated as a flat fading channel with constant channel gain. Provided that the system parameters are accurately dimensioned, OFDM transmissions are not affected by ISI even in highly frequency-selective channels [4].

Several studies have shown that in a single-user system large gains in terms of system performance are given by employing dynamic resource allocation algorithms, as compared to static allocation schemes. The goal is to select the best set of modulation parameters for each subcarrier so as either to maximize the overall data throughput under a constraint on total transmitted power or to minimize the overall transmit power given a fixed throughput.

In a multiuser scenario, the problem of resource allocation arises from the need for a multiple access scheme in which users share the same bandwidth. In static allocation schemes, each user receives predetermined time slots or frequency channels respectively without considering their channel conditions. The problem reduces to only power allocation on the subcarriers. However, since the fading parameters for different users are mutually independent, the probability that a subcarrier is in deep fade for all users is very low. Thus, in an OFDMA system where each user is assigned a different subset of the available bandwidth, assuming that the transmitter has perfect knowledge of CSI for each user, subcarriers can be assigned according to a predetermined optimization criterion so as to increase the system spectral efficiency.

Most of the existing literature focuses on the single-cell scenario, either formulating the problem with the goal of maximizing the system sum rate subject to a limitation on power consumption, or aiming at minimizing the overall transmit power subject to users' rate constraints. Only recently, the problem of resource allocation in multi-cell networks has been addressed. In multi-cell environments, resource allocation is also helpful in reducing MAI from neighboring cells, thus making possible the development of a cluster with full reuse of the frequency spec-

trum. Unfortunately, the complexity of the allocation problem is extremely demanding and also, due to the detrimental effect of the MAI, the solution may not exist.

Among multicarrier modulation techniques, a valid alternative to conventional OFDM signaling is represented by FBMC modulations initially proposed for very high-speed wired access networks. In FBMC transmissions data symbols are frequency-multiplexed over contiguous subchannels after proper pulse shaping [5]. Compared to OFDM modulation, pulses are significantly longer than the subchannel symbol spacing, and thus overlap in time. Conversely, signal spectra on the subcarriers are band-limited and, depending on the shaping filter employed in the transmission filter bank, they can be either non-overlapping in frequency or marginally overlapping as in [6]. As a result, FBMC can be envisaged as an efficient alternative to conventional OFDM with a number of attractive features, such as *i*) lesser sensitivity to narrowband interferers, *ii*) higher flexibility to allocate groups of subchannels to different users, *iii*) mitigation of ICI on severely time-frequency selective channels, *iv*) simpler frequency domain equalization not requiring any cyclic extension. The above qualities explain why FBMC schemes have been adopted as well for a number of wireless standards, such as the return channel of terrestrial DVB (DVB-RCT) [7] and the 2^{nd} release of the TETRA air interface [8].

Like any other multicarrier transmission schemes, FBMC is strongly sensitive to residual CFOs, which must be accurately estimated and removed from the received waveform prior to channel estimation and data decoding. The issue of CFO recovery for multicarrier systems has received considerable attention in recent literature, but unfortunately the proposed techniques are primarily intended for OFDM and are neither optimized nor directly applicable to FBMC in view of the rather different signal formats. Until now, only a few algorithms for CFO extraction specifically tailored for FBMC have been proposed. In addition, they are devised under the assumption of time-invariant or slowly changing fading and as such they exhibit poor behavior whenever the channel is affected by a significant Doppler spread.

In this dissertation we also deal with FBMC communications and we propose a novel CFO recovery algorithm that, rather than either relying on known pilot symbols multiplexed within the transmitted burst or exploiting the specific signal structure in a blind mode, takes advantage of both pilot symbols and also indirectly of the data symbols through differential decisions. We will show that, when compared to conventional pilot-based methods, the proposed approach improves the frequency acquisition range without degrading accuracy, while retaining approximately the same computational load.

1.1 Outline of Dissertation

The outline of the dissertation is as follows.

The present Chapter introduces the motivations and summarizes the structure of this dissertation.

Chapter 2 provides the basic concepts of OFDM and FBMC systems.

In Chapter 3, in order to introduce and discuss the main contribution of this dissertation, we briefly revise the well-known resource allocation algorithms for single-cell OFDMA systems.

Chapter 4 analyzes and discusses the performance of a resource allocation scheme for an OFDMA multi-cell system. We design a layered architecture that integrates a packet scheduler and an adaptive resource allocator with the goal of mitigating the multiple access interference. This architecture is able to assign radio channels to the users so as to minimize interference, as well as to guarantee a fair number of resources to each user in the system.

In Chapter 5 we deal with FBMC systems and develop an algorithm for the recovery of the CFO over time-frequency selective fading channels. The algorithm we derive is based on the ML principle and takes advantages of both pilot symbols and also indirectly of the data symbols through differential decisions.

Chapter 2

Multicarrier modulation techniques

We pointed out that in frequency-selective fading channels the received signal is typically affected by ISI. The classical approach adopted in single carrier systems is time-domain equalization. However, the number of operations per signaling interval grows linearly with the number of interfering symbols or, equivalently, with the data rate. As a result, conventional time-domain equalizers are not suitable for high-speed transmissions with channel delay spreads extending over tens of symbol intervals. This has motivated the adoption of multicarrier modulations as a computationally efficient alternative for facing with the severe impairments of multipath propagation. The idea behind multicarrier modulations is to split a high-rate data stream into a number of low-rate streams which are transmitted in parallel on adjacent subchannels. As is shown in Figure 2.1, reducing the data rate or, equivalently, increasing the symbol duration, makes the frequency selective fading channel appear flat on each subcarrier, thereby limiting the ISI.

In the next, we briefly revise the system structure of these two multicarrier technologies, such as OFDM and FBMC, which have been adopted by several standards for next generation communication systems.

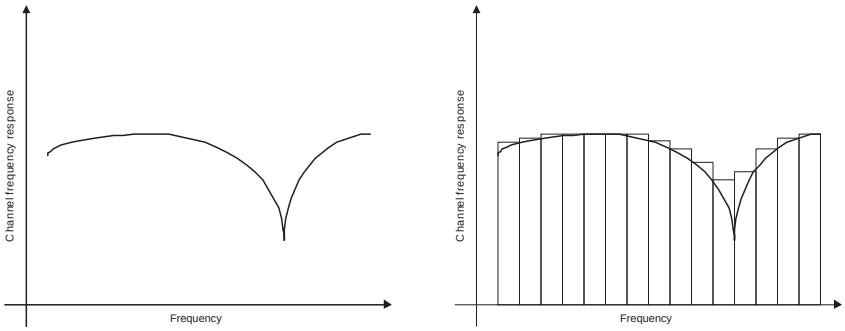

Figure 2.1: Channel frequency response. In multicarrier system each information-bearing symbol undergoes frequency flat fading channel.

2.1 OFDM

OFDM is a multiplexing technique in which a high data rate stream is split into N subflows with lower data rate and the N OFDM symbols are transmitted simultaneously over the N orthogonal subcarriers. Compared with single-carrier multiple-access systems, OFDM offers increased robustness to narrowband interference and does not need adaptive time-domain equalizers, since channel equalization is performed in the frequency domain through one-tap multipliers. The combination of OFDM with a dynamic channel assignment algorithm allows the system to increase its performance in terms of capacity.

This technique was originally implemented using a bank of Nyquist filters which provide a set of continuous-time orthogonal basis functions. Using very fast and cost effective digital signal processors, OFDM modulation is now implemented using DFT techniques. This has motivated the adoption of OFDMA as a standard for the WLAN IEEE 802.11a, Wi-MAX 802.16 and it has also been proposed for digital cable television systems.

2.1.1 Transmitter structure

The block diagram of the transmitter is shown in Figure 2.2. The complex symbols c_i belonging to an M-QAM or M-PSK constellation with average energy σ_c^2 are characterized by a signaling rate $R = 1/T$. These symbols are fed to a serial-to-parallel converter, where each flow has a signaling time $T_s = NT$, which is referred to as OFDM symbol time. In the figure, the index n ($n = 1, \ldots, N$) denotes the subcarrier, while m is the index for an OFDM symbol with rate $1/T_s = (1/T)/N$. Thus, the N symbols at the output of the serial-to-parallel converter are

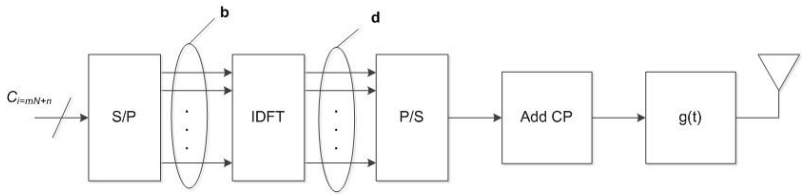

Figure 2.2: Block diagram of an OFDM transmitter.

sent to an OFDM modulator, which comprises an N-point IDFT unit and the insertion of an N_G-point CP larger than the channel impulse response. The CP serves to eliminate inter-block interference and makes the linear convolution of the symbols with the channel look like a circular convolution, which is essential for demodulation based on DFT. This produces an $(N + N_G)$-dimensional vector $\mathbf{d} = [d(-N_G), d(-N_G + 1), \ldots, d(N - 1)]^T$ of time domain samples where $d(n) = d(n + N)$ for $-N_G \leq n \leq -1$ and

$$d(n) = \sum_{\ell=1}^{N} b(\ell) e^{-j2\pi \frac{\ell n}{N}}, \quad 0 \leq n \leq N - 1, \tag{2.1}$$

where $b_\ell = c_0(\ell)$ are the transmitted symbols on the first OFDM symbol.

The resulting vector $\mathbf{d}$ is finally passed to a D/A converter with impulse response $g(t)$ and signalling interval T_s. The complex envelope of the signal transmitted takes the form

$$s(t) = \sum_{n=-N_G}^{N-1} d(n) g(t - nT_s) \tag{2.2}$$

where $g(t)$ has a root-raised cosine Fourier transform with some roll-off α.

2.1.2 Receiver structure

Figure 2.3 illustrates the block diagram of an OFDM receiver. At the receiver, the incoming waveform is first filtered and then sampled at rate $1/T_s = 1/(NT)$. This produces

$$x(n) = \sum_{\ell=-N_G}^{N-1} d(\ell) h(n - \ell) + w(n) \tag{2.3}$$

where $w(n)$ is thermal noise and $h(n)$ is the sample of the overall channel impulse response $h(t)$ (including the physical channel, the shaping pulse at the transmitter and the whitening matched filter at the receiver) at time $t = nT_s$. Statistical models for the channel impulse response of a fading multipath channel have been described in details in literature over the past

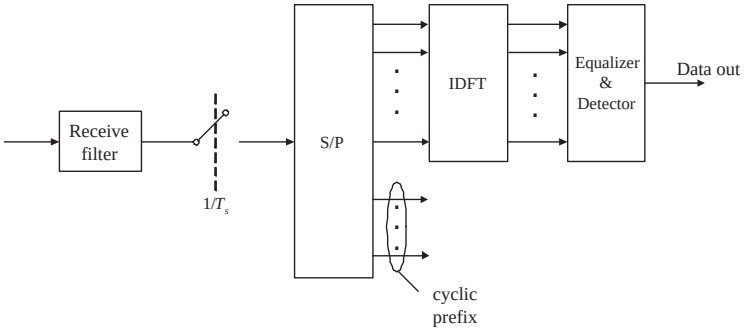

Figure 2.3: Block diagram of an OFDM receiver.

years. Since its out of the scope of this work to provide a description of such models, we refer to some excellent references on this field such as [9]-[10].

After discarding the CP and using a matrix notation, the samples at the output of the matched filter can be expressed as follows

$$\begin{bmatrix} x(0) \\ x(1) \\ \vdots \\ x(N-1) \end{bmatrix} = \begin{bmatrix} h(0) & 0 & \cdots & \cdots & 0 & h(L) & \cdots & h(1) \\ \vdots & \ddots & \ddots & & & 0 & \ddots & \vdots \\ \vdots & & h(0) & \ddots & & & \ddots & h(L) \\ h(L) & & \vdots & \ddots & 0 & & & 0 \\ 0 & \ddots & \vdots & & h(0) & \ddots & & \vdots \\ \vdots & \ddots & h(L) & & \vdots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \vdots & & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & h(L) & \cdots & \cdots & h(0) \end{bmatrix} \begin{bmatrix} d(0) \\ d(1) \\ \vdots \\ d(N-1) \end{bmatrix} \tag{2.4}$$

where L is the length in sampling periods of the channel impulse response. Inspections of the above equation reveals that thanks to the introduction of the CP at the transmitter and its removal at the receiver, the resulting time domain samples are related to the input data symbols through the channel matrix in (2.4). The latter is a circulant matrix, i.e., its rows are composed of cyclically shifted versions of a given sequence [11]. In other words, the effect of the cyclic prefix is to make the channel look like circular convolution instead of linear convolution, thereby completely removing the ISI.

From matrix theory it turns out that these kind of matrices have a very interesting and useful property [12]. The eigenvectors are independent of the specific channel coefficients and are always given by complex exponentials. To be more precise, the EVD of the circulant channel

matrix in (2.4) is

$$\begin{bmatrix} h(0) & 0 & \cdots & \cdots & 0 & h(L) & \cdots & h(1) \\ \vdots & \ddots & \ddots & & & 0 & \ddots & \vdots \\ \vdots & & h(0) & \ddots & & & \ddots & h(L) \\ h(L) & & \vdots & \ddots & 0 & & & 0 \\ 0 & \ddots & \vdots & & h(0) & \ddots & & \vdots \\ \vdots & \ddots & h(L) & & \vdots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \vdots & & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & h(L) & \cdots & \cdots & h(0) \end{bmatrix} = \mathbf{F}^H \mathbf{H} \mathbf{F} \tag{2.5}$$

where $\mathbf{F}$ is the $N \times N$ unitary DFT matrix whose entries are given by

$$[\mathbf{F}]_{n,\ell} = e^{-j2\pi\, n\ell/N} \qquad 0 \le n \le N-1\,, \quad 0 \le \ell \le N-1. \tag{2.6}$$

while $\mathbf{H} = \mathrm{diag}\,\{\mathrm{H}(0), \mathrm{H}(1), \ldots, \mathrm{H}(\mathrm{N}-1)\}$ is $N \times N$ diagonal matrix with

$$H(n) = \sum_{\ell=0}^{L-1} h(\ell) e^{-j2\pi \frac{\ell n}{N}}. \tag{2.7}$$

Collecting the above fact together, we see that at the receiver the DFT outputs $y(n)$ for $n = 0, 1, \ldots, N-1$ can be written as

$$y(n) = H(n)s(n) + w(n) \tag{2.8}$$

or in matrix notation

$$\mathbf{y} = \mathbf{H}\mathbf{s} + \mathbf{w}. \tag{2.9}$$

From the above equation it follows that thanks to the multicarrier approach, the original frequency-selective channel with inter-symbol and inter-block interference is transformed into a set of parallel flat subchannels, which can be straightforwardly equalized at receiver side using a simple bank of one-tap multipliers.

2.2 FBMC

2.2.1 Transmitter structure

In FBMC-based systems, the data symbols are transmitted over different subchannels after suitable pulse shaping. Unlike OFDM, the pulse waveforms associated to different symbols

overlap in time, As a result of pulse shaping, the spectra of the data signals on the different subcarriers are bandlimited, while preserving orthogonality amongst subcarriers in spite of time overlap.

Let us assume that the source M-QAM data symbols $a_\ell^{(q)}$ (at the rate N/T) are grouped into blocks of size N, the symbol q being the index of the symbol within each block ($1 \leq q \leq N$) and is also the subcarrier index after MC modulation. Each subcarrier is shaped by means of a conventional SRRC filter with impulse response $g(t)$ and roll-off ξ and a signaling interval T. In order to prevent spectral overlap, each subcarrier is centered at frequency $q(1+\xi)/T$. Assuming that the quantity $M = (1+\xi)N$ is an integer, the transmitted signal can be expressed as

$$s(t) = \sum_{\ell} \sum_{q=1}^{N} a_\ell^{(q)} g(t - \ell T) e^{j2\pi q M t/NT}. \tag{2.10}$$

In a digital implementation of the modulator using the sampling frequency $f_s = 1/T_s = M/T$, the digitized FBMC signal is thus

$$s_h^{(m)} \triangleq s((Mh+m)T_s) = \sum_{\ell} \sum_{q=1}^{N} a_\ell^{(q)} g[M(h-\ell)+m] e^{j2\pi q(Mh+\ell)/N}, \tag{2.11}$$

where we adopt a polyphase decomposition for the time index of the i-th sampling instant $t_i = (Mh+m)T_s$, $1 \leq m \leq M$ and h some integer, and $g[\ell] \triangleq g(\ell T_s)$.

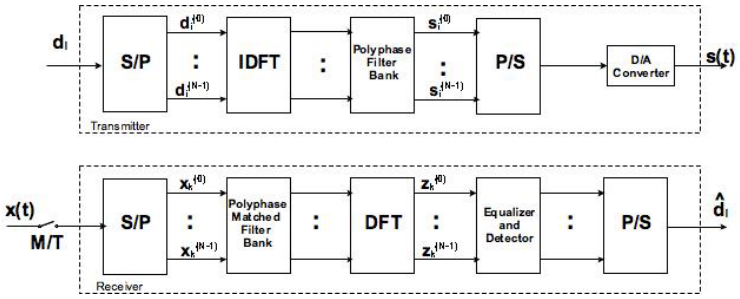

Figure 2.4: Block diagram of an FBMC transmission systems.

Figure 2.4 shows the FBMC transmission system, which comprises an OFDM modulator followed by a bank of polyphase filter. The ℓ-th N symbol block are fed to an IDFT, where each element of the parallel IDFT output is processed at the rate $1/T$ by a different filter (the *filterbank*) whose impulse response $g_T^{(m)}[\ell]$ is obtained by the polyphase decomposition of the prototype SRRC filter

$$g_T^{(m)}[\ell] = g\left(\ell T + \frac{mT}{M}\right) = g(\ell M T_s + m T_s), \quad 1 \leq m \leq M. \tag{2.12}$$

The samples obtained at the output of the filter bank are then rearranged and D/A converted to produce the time-continuous FBMC signal (2.10). It is worth pointing out that the number of filters is M (with M being the number of signal samples to be computed in every block of N source symbols), while the number of IDFT outputs is N (i.e., the number of subcarriers). Thereby, the filterbank $g_T^{(m)}[\ell]$ also contains a suitable signal memory that feeds the diverse filters of the bank with appropriately permuted versions of the IDFT outputs [5].

2.2.2 Receiver structure

Let us assume that the propagation channel is selective in time and frequency with impulse response

$$c(t) = \sum_{u=1}^{U} \rho_u(t)\delta(t-\tau_u), \tag{2.13}$$

where U is the number of paths, while $\rho_u(t)$ and τ_u are the time-varying complex gains and the delays of each path, respectively. Then, the received signal can be expressed by

$$\begin{aligned} r(t) = \sum_{u=1}^{U}\sum_{\ell}\sum_{q=1}^{N} \rho_u(t) a_\ell^{(q)} g(t-\ell T-\tau_u) e^{j2\pi qMt/NT} \\ \cdot\, e^{-j2\pi qM\tau_u/NT} e^{j2\pi\nu t/T} e^{j\theta} + n(t), \end{aligned} \tag{2.14}$$

where ν is the frequency offset normalized to the signaling rate $1/T$, θ is the phase offset and $n(t)$ is the AWGN with two-sided power spectral density $2N_0$.

In the lower part of Figure 2.4 it is shown the implementation of FBMC receiver. The samples are taken at rate M/T and processed by the receiver polyphase filterbank (i.e., receiver matched filter) and the DFT unit. Then, data detection is performed after channel equalization [5], [13]. The output of the DFT corresponding to the n-th subcarrier at the instant $t = kT$ is given by

$$z_k^{(n)} = \sum_{u=1}^{U}\sum_{\ell}\sum_{q=1}^{N} a_\ell^{(q)} A_{k,\ell,u}^{(n,q)}(\nu) e^{j\theta} + w_k^{(n)}, \tag{2.15}$$

where $w_k^{(n)}$ is a zero-mean independent Gaussian random variable with variance $\sigma^2 = 2N_0$ and

$$\begin{aligned} A_{k,\ell,u}^{(n,q)}(\nu) \triangleq e^{-j2\pi qM\tau_u/NT} \int_{-\infty}^{+\infty} \rho_u(t) g(t-\ell T-\tau_u) \\ \cdot\, g(t-kT) e^{j2\pi(q-n)Mt/NT} e^{j2\pi\nu t/T} dt \end{aligned} \tag{2.16}$$

is a complex-valued factor accounting for the contribution to the sample (2.15) of the ℓ-th symbol on the q-th subcarrier propagating over the u-th path in the presence of the frequency offset ν.

Assume that within a time Δt comparable with the non-zero support of $g(t)$ (namely, a few FBMC symbol intervals) the following assumptions hold: *i*) the fading process does not change significantly; *ii*) the frequency offset is small enough so that $e^{j2\pi\Delta t/T} \approx 1$; *iii*) the delay spread of the channel is significantly shorter than the FBMC signaling interval. More in detail, assumptions *i*) and *ii*) allow us to consider as orthogonal the signals on different subchannels, while assumption *iii*) implies that the channel delay spread does not destroy the orthogonality of the consecutive symbols on the same subcarrier, i.e., the ISI on each subcarrier is negligible.

Based on this discussion, we can simplify the (2.16) as follows

$$A_{k,\ell,u}^{(n,q)}(\nu) \simeq \begin{cases} \rho_u(kt)e^{-j2\pi qM\tau_u/NT}e^{j2\pi\nu t/T}, & q=n,\ \ell=k \\ 0, & \text{otherwise} \end{cases} \tag{2.17}$$

thereby, (2.15) can be approximated as

$$z_k^{(n)} \simeq \varphi_k^{(n)} a_k^{(n)} e^{j2\pi\nu k} e^{j\theta} + w_k^{(n)}, \tag{2.18}$$

where

$$\varphi_k^{(n)} = \sum_{u=1}^{U} \rho_u(kT)e^{-j2\pi nM\tau_u/NT} \tag{2.19}$$

is a multiplicative factor accounting for the channel time-frequency selectivity. Notice that the expression (2.18), even in presence of small frequency offsets, is different from the one obtained in the context of OFDM. Indeed, due to the overlapping of subcarrier spectra, the OFDM model would necessarily include ICI from *all* subcarriers. On the other hand, in FBMC transmissions, ICI is absent thanks to the subcarrier separation, although large frequency offset or large Doppler spreads (a condition that is rejected by assumption *ii*)) might lead to considerable ICI between neighboring spectra.

Chapter 3

Resource allocation in single cell OFDMA systems

In future multicarrier systems the communications are characterized by ever higher data rates and the bandwidth requirements of modern equipment are constantly increasing. We already pointed out that one of the main limitation of wireless communications is represented by the frequency-selective nature of the wireless channel, so that more and more wireless devices employs OFDM signaling to reduce the impact of the ISI on the received signal. Moreover, when OFDM is adopted along with adaptive resource allocation algorithms, it is possible to increase the system spectral efficiency by exploiting the system frequency and multi-user diversity.

In this Chapter we analyze and evaluate the performance of several intelligent dynamic resource allocation schemes for OFDM-based systems, focusing on both single- and multi-user OFDMA scenario. In single-user OFDM systems adaptive channel assignment schemes are derived with the goal of taking advantage of the frequency selectivity of the channel by choosing the best set of modulation type and transmit power for each subcarrier [14]- [17]. Two main approaches can be considered, depending on whether the system tries to maximize the overall data rate with a constraint on power consumption or to minimize the overall transmit power subject to user's rate constraints. The problem of adaptively allocating information bits on each subcarrier is defined as *bit loading*. Any bit loading algorithm tends to transmit more information on those channels exhibiting better conditions, i.e., having the highest SNRs. On the contrary, small-size constellations are employed on subcarriers affected by severe fading. The problem of adjusting transmit power levels on each subcarrier is referred to as *power*

allocation. We will see that by applying the water-filling algorithm, it is possible to achieve the theoretical capacity in a frequency-selective channel.

In a multi-user OFDM system, due to the different spatial positions occupied by each terminal, signals received by each user undergo independent fading attenuations. Thus, a subcarrier that appears in a deep fade for one terminal may have a much higher gain for other users. If the transmitter has perfect knowledge of CSI for each user, subcarriers and transmit power are dynamically assigned with respect to some predetermined optimization criterion, so that the system spectral efficiency can be increased by exploiting the so-called *multi-user diversity*.

With respect to single-user systems, optimum resource allocation in a multiuser scenario involves the use of subcarrier assignment scheme along with adaptive power and bit allocation algorithms. This leads to a more challenging task than in a single-user system. Provided that users cannot utilize the same channel, the allocation problem turns to be a combinatorial optimization problem for which no optimal greedy solution exists. Several research activities have been carried out in literature to develop suboptimal resource allocation strategies with good performance and limited complexity.

In this Chapter we present and discuss different strategies for resource allocation in single cell OFDM system. We start analyzing the bit and power loading for a single-user scenario. We revisit the classical water-filling principle in order to maximize the capacity of the user and discuss practical schemes based on rate maximization or transmit power minimization. After that, we present several examples of dynamic resource allocation algorithms for multi-user OFDM systems, where rate maximization or power minimization are extended to a typical OFDMA scenario. We describe the optimal joint channel and power allocation, whose computational complexity is too large, then we derive suboptimal schemes with limited complexity.

3.1 Resource allocation in single-user OFDM systems

In single-user OFDM systems, the resource allocation problem consists of determining an appropriate transmit power allocation across the subcarriers. Based on the model described in Chapter 2, assuming perfect timing and frequency synchronization, the output of the receive DFT can be expressed as

$$y(n) = H_n s(n) + w(n) \quad n = 1, \ldots, N \tag{3.1}$$

where H_n is the channel frequency response on the n-th subchannel, $s(n)$ is the corresponding input symbol with power $P_n = E\{|s(n)|^2\}$ and $w(n)$ is the zero-mean AWGN with two-sided power spectral density equal to $2N_0$. We use Shannon capacity [18] as the primary measure of the rate achievable by a certain user on a particular channel. Given a certain SNR, the channel capacity in bit/s/Hz is $\eta = \log_2(1 + SNR)$ and the maximum theoretical rate in bit/s achievable over a channel that spans the bandwidth B is $R = B\eta$. Thus, the data rate in bit/s for subcarrier n $(n = 1, \ldots, N)$ is given by

$$r_n = B \log_2(1 + \gamma_n), \tag{3.2}$$

where B is the bandwidth of a subcarrier and γ_n is the SNR of the n-th subcarrier given by

$$\gamma_n = \frac{P_n |H_n|^2}{\sigma_w^2}, \tag{3.3}$$

with $\sigma_w^2 = BN_0$. Equation (3.2) is the data rate achieved with an arbitrarily low error probability. However, practical communication systems are normally designed for a non-zero target BER which determines the requested QoS. We define the *SNR gap* as the difference between the SNR necessary to achieve a certain data rate for real systems and the theoretical limit. For example, the BER for an AWGN channel with M-QAM modulation and ideal phase detection is upper bounded by [19]:

$$P_e \leq 2e^{-1.5\gamma/(M-1)}, \tag{3.4}$$

where $M = 2^\eta$, η is the spectral efficiency in bit/s/Hz, while γ is the SNR as defined in (3.3). When $\eta \geq 2$ and $0 \leq \gamma \leq 30$ dB, BER could be approximated within 1dB by [20]:

$$P_e \leq 0.2e^{-1.5\gamma/(M-1)}. \tag{3.5}$$

Using (3.5), the rate for subcarrier n is given by:

$$r_n = B \log_2(1 + \frac{\gamma_n}{\Gamma_n}), \tag{3.6}$$

where Γ_n is the SNR gap that can be expressed as:

$$\Gamma_n = -\frac{\ln(5P_{e,n})}{1.5}, \tag{3.7}$$

$P_{e,n}$ denoting the BER on subcarrier n. Note that when is $\Gamma_n = 1$, it is $\gamma_n = 2^\eta - 1$ and the rate on subcarrier n is given by the Shannon formula, i.e., $r_n = B\eta$. Depending on the value of BER desired on each subcarrier, the effective SNR has to be adjusted according to the modulation scheme. For example, power and data rate are allocated on the subcarriers so as

the BER across each tones does not exceed a given threshold [21], [15], [22]. Another strategy is to specify an average BER over the OFDM block, thus resulting into a non-uniform BER across subcarriers [23], [24].

Regardless of the policy adopted on BER, resource allocation algorithms can be divided into RA algorithms and MA algorithms. The RA schemes aim at maximizing the user data rate under a limitation on transmit power, namely

$$\max_{\{P_n\}} R_T = \sum_{n=1}^{N} r_n, \tag{3.8}$$

$$\sum_{n=1}^{N} P_n = P_{tot}. \tag{3.8.1}$$

On the other hand, following the MA approach we allocate bit and power with the goal of minimizing the overall transmit power given a fixed data rate constraint:

$$\min_{\{P_n\}} P_T = \sum_{n=1}^{N} P_n \tag{3.9}$$

$$\sum_{n=1}^{N} r_n = R_{\text{target}}. \tag{3.9.1}$$

3.1.1 The water-filling algorithm

We first consider the RA problem as defined in (3.8) and derive its optimal solution based on the water-filling algorithm. The optimization problem in (3.8) is convex, since both the objective function and the constraints are convex, thus the optimal solution can be found using Lagrangian dual decomposition [25]. The Lagrangian of problem (3.8) is

$$\mathcal{L} = B \sum_{n=1}^{N} \log_2 \left(1 + \frac{P_n \left| H_n \right|^2}{\sigma_w^2} \right) + \lambda \left(P_{tot} - \sum_{n=1}^{N} P_n \right), \tag{3.10}$$

where λ is the Lagrangian multiplier. By deriving the Lagrangian $\mathcal{L}$ with respect to P_n we obtain

$$\frac{\partial \mathcal{L}}{\partial P_n} = \frac{B}{(P_n + \sigma_w^2/|H_n|^2)\log 2} - \lambda. \tag{3.11}$$

The optimal power allocation satisfying (3.11) is

$$P_n^{(opt)} = \left(\mu - \frac{1}{g_n} \right)^+, \tag{3.12}$$

where $g_n \triangleq |H_n|^2/\sigma_w^2$ is the so-called *channel SNR*. Also, $\mu = B/(\lambda \log 2)$ is a parameter that must be chosen so as to satisfy the total transmit power constraint

$$\sum_{n=1}^{N} \left(\mu - \frac{1}{g_n} \right)^+ = P_{tot}. \tag{3.13}$$

The optimal solution of (3.8) is also illustrated in Figure 3.1 and gives an interesting physical interpretation. Indeed, the quantities can be seen as the bottom of a vessel in which the transmit power P_{tot} is poured as it were water. The level of water is represented by the quantity μ, while $P_n^{(opt)}$ is the amount of water on subcarrier n. Also, when the bottom level for a given subcarrier is higher than the water surface, then no power is allocated on the corresponding subcarrier since it is too faded to support reliable data transmission. The idea behind the water-filling approach is to give more power to those channels with high quality (i.e., those with high SNR), while the subcarriers characterized by low SNR receive less power or even are left unused.

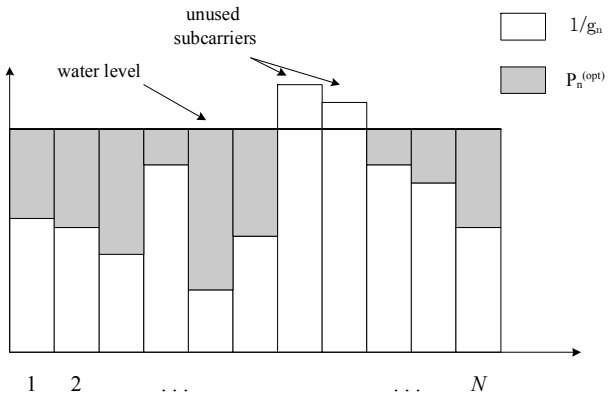

Figure 3.1: Water-filling

Unfortunately, due to the non-linear relationship between the total power constraint and the quantities $1/g_n$ and μ, the optimal power allocation given by (3.12) can only be found by means of an iterative procedure where the quantity μ is recalculated at each new iteration after eliminating those subcarriers with the lowest channel SNRs. In detail, let $\mathcal{N}^{(i)}$ be the set of subcarrier that are considered during the i-th iteration, with $\mathcal{N}^{(0)} = \{1, 2, \ldots, N\}$. Then, the water level is first calculated by (3.13) as

$$\mu^{(i)} = \frac{1}{\text{card}\{\mathcal{N}^{(i)}\}} \left(P_{tot} + \sum_{n \in \mathcal{N}^{(i)}} \frac{1}{\gamma_n} \right). \tag{3.14}$$

This value is then substituted into (3.12) to obtain the tentative power allocated on the n-th subcarrier

$$P_n^{(i)} = \begin{cases} \mu^{(i)} - 1/g_n & \text{if } n \in \mathcal{N}^{(i)} \\ 0 & \text{otherwise} \end{cases} \tag{3.15}$$

At the end of each iteration, subcarriers with negative power assignments are discarded from the set $\mathcal{N}^{(i)}$ and their power value set to zero. For each new iteration, only those subcarriers

with power greater than zero are considered. The algorithm ends when all power values P_n are non-negative.

We now focus on the MA approach, as stated in (3.9), and we derive its optimal solution similarly to what we have done for the RA case. The Lagrangian of problem (3.9) is

$$\mathcal{L} = \sum_{n=1}^{N} P_n + \lambda \left(R_{\text{target}} - B \sum_{n=1}^{N} \log_2 \left(1 + \frac{P_n |H_n|^2}{\sigma_w^2} \right) \right) \tag{3.16}$$

where λ is the Lagrangian multiplier. The derivative of $\mathcal{L}$ with respect to P_n is

$$\frac{\partial \mathcal{L}}{\partial P_n} = 1 - \lambda \frac{1}{1+\gamma_n} \cdot \frac{|H_n|^2}{\sigma_w^2} \cdot \frac{1}{\log 2}. \tag{3.17}$$

Thus, the optimal power value on subcarrier n is given by

$$P_n^{(opt)} = \left(\mu - \frac{1}{g_n} \right)^+ \tag{3.18}$$

where $\mu = \lambda B / \log 2$ is the water level such that

$$B \sum_{n=1}^{N} \log_2 (\mu g_n) = R_{\text{target}}. \tag{3.19}$$

As described for the water-filling in the RA case, μ can be seen as the common water level of the power or water that is poured over channels, with river beds being equal to $1/g_n$. The algorithm is iterative and starts with the maximum number of streams, then μ is updated for a decreasing number of streams until the point where the water level is above the highest river bed.

3.2 Multi-user OFDM systems

The classification of resource allocation schemes that has been carried out in Section 3.1 for single-user OFDM networks can also be utilized for multi-user OFDM scenarios. The goal of RA algorithms is to maximize the system data rate with the constraint on the total transmit power, while MA algorithms aims at minimizing the total transmit power subject to rate constraints per user [26]-[28]. Moreover, RA schemes are divided into two major groups based on the user constraints. In the first group, there is a fixed rate requirement for each user and the goal is to maximize the total throughput of the system while satisfying individual users' rate requirements [29]-[31] . In the second group [32]-[34] the concept of fairness is utilized. Here the objective is not only to maximize the total throughput subject to a total transmit power constraint, but

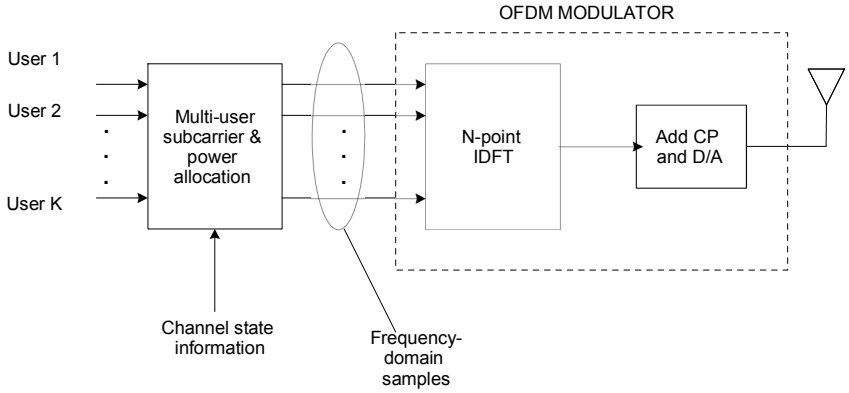

Figure 3.2: BS transmitter for OFDMA downlink transmission with adaptive resource allocation

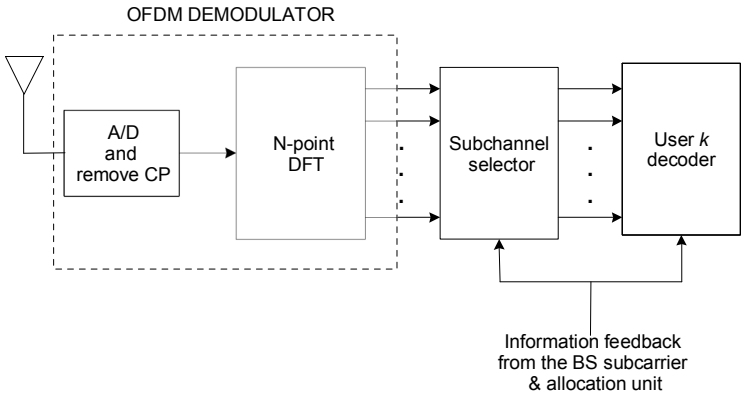

Figure 3.3: k-th user receiver for OFDMA downlink transmission with adaptive resource allocation

it is also to maintain the rate proportionality among the users based on given proportional constraints rather than achieving a specific requested data rate.

The concept of dynamic resource allocation in an OFDMA downlink transmission is illustrated in Figure 3.2 and Figure 3.3. At the BS, the information about the users' channel responses are sent to the multi-user subcarrier and power allocation unit, which maps the users' data over the selected subcarriers choosing the best transmission mode (coding and/or modulation scheme). The complex symbols at the output of the modulators are fed to an OFDM modulator and transmitted over the channel.

At the k-th user, the received signal is demodulated and the recovered frequency-domain samples are passed to subchannel selector, which takes into account the information about

channel allocation of user k. Then, the selected samples are decoded, thus providing the final detected bit by means of an appropriate detection strategy. The information about subcarrier and power allocation are sent by the BS to each user on a control channel. This exchange of information leads to a certain transmission overhead which tends to reduce the achievable data throughput. This problem can be mitigated by grouping sets of adjacent subcarriers into *subchannels* with similar fading characteristic. As a result, assuming that the bandwidth of a subchannel is smaller than the channel coherence bandwidth, the channel spectrum can be approximated as flat in the sub-channel. Resource allocation is thus performed on subchannels rather than on subcarriers causes almost no loss in diversity. These assumptions are coherent with that is being done in existing OFDMA systems. For example, in the WiMAX [2] band AMC zone contiguous subcarriers and OFDM symbols are grouped together into one basic allocation unit, called a *sub-channel*. Similarly, in the UTRAN LTE system [3] the subcarriers in each TTI are grouped into equal sized *physical resource blocks*; each block is allocated to a single user, and a user can be scheduled on multiple blocks. In the remainder, unless differently stated, we have neglected the impact of the control feedback channel on the system performance.

We consider a downlink communication in an OFDMA system where the BS and K users are equipped with a single antenna. The overall frequency spectrum is divided into N orthogonal subcarriers. We also assume that the BS has perfect knowledge of CSI for each user, thus allowing the BS to dynamically allocate power and subcarriers according to channel quality. Using the Shannon capacity as a measure of data-rate achieved on a given subchannel, the data rate for user k is

$$R_k = \sum_{n=1}^{N} r_{k,n} = B \sum_{n=1}^{N} a_{k,n} \log_2(1 + \gamma_{k,n}), \tag{3.20}$$

where $a_{k,n}$ is a binary variable that assumes value 1 if subcarrier n is allocated to user k, and zero otherwise. $\gamma_{k,n}$ is the SNR for user k on subcarrier n, defined as

$$\gamma_{k,n} = \frac{P_n H_{k,n}}{\sigma_w^2}, \tag{3.21}$$

with $H_{k,n}$ denoting the channel gain (accounting for pathloss and multipath fading) between user k and BS on the n-th link.

3.2.1 Multi-user RA algorithms

The RA resource allocation problem is formulated as follows:

$$\max_{\{a_{k,n}\},\{P_n\}} \sum_{k=1}^{K}\sum_{n=1}^{N} r_{k,n} \tag{3.22}$$

$$\sum_{k=1}^{K} a_{k,n} \leq 1 \quad n = 1,\ldots,N \tag{3.22.1}$$

$$\sum_{n=1}^{N} P_n \leq P_{tot}, \tag{3.22.2}$$

$$a_{k,n} \in \{0,1\} \quad k = 1,\ldots,K \quad n = 1,\ldots,N \tag{3.22.3}$$

Constraints are described in (3.22.1)-(3.22.3): (3.22.1) implements orthogonal assignment, i.e., one subcarrier can be allocated to at most one user and, while in equation (3.22.2) P_{tot} is the maximum allowable transmit power.

Problem (3.22) is not convex since it needs to find the optimal set of subcarriers for each user, which turns out to be a combinatorial problem whose complexity increases exponentially with N. Indeed, finding optimal channel assignment requires K^N searches, thus the overall optimization involves $\mathcal{O}(NK^N)$ operations. In detail, the k'-th user is given the n-th subcarrier on condition that

$$k' = \arg \max_{1\leq k\leq K} \{\gamma_{k,n}\}. \tag{3.23}$$

while the power allocated on subcarrier n is given by

$$P_n = \left(\mu - \frac{1}{g_{k',n}}\right)^+ \tag{3.24}$$

with $g_{k',n} \triangleq |H_{k',n}|^2/\sigma_w^2$ is the corresponding channel SNR for user k' on subcarrier n and μ is a parameter that have to be set so as to satisfy the following condition

$$\sum_{n=1}^{N} \left(\mu - \frac{1}{g_{k',n}}\right)^+ = P_{tot}. \tag{3.25}$$

The optimal solution to the non-convex optimization problem (3.22) can be achieved by means of Lagrangian dual function. It has been shown [35] that in multicarrier applications, even though the original resource allocation problems are non-convex, the duality gap becomes zero as the number of subchannels goes to infinity.

The Lagrangian of problem (3.22) is defined over domain $\mathcal{D}$ as

$$\mathcal{L}(\{P_n\},\{r_{k,n}\},\lambda) = \sum_{k=1}^{K}\sum_{n=1}^{N} r_{k,n} - \lambda\left(\sum_{n=1}^{N} P_n - P_{tot}\right), \tag{3.26}$$

where $\mathcal{D}$ is defined as the set of all $P_n > 0$ for $k = 1,\ldots,K$ and $n = 1,\ldots,N$. The Lagrange dual function is

$$g(\lambda) = \max_{\{P_n\},\{r_{k,n}\}\in\mathcal{D}} \mathcal{L}(\{P_n\},\{r_{k,n}\},\lambda). \tag{3.27}$$

Equation (3.26) suggests that the maximization of $\mathcal{L}(\{P_n\},\{r_{k,n}\},\lambda)$ can be decomposed into N independent optimization problems,

$$g_n^{'}(\lambda) = \max_{\{P_n\}\in\mathcal{D}} \left\{\sum_{k=1}^{K} r_{k,n} - \lambda P_n\right\} \quad n = 1,\ldots,N \tag{3.28}$$

Thus, the Lagrange dual function becomes

$$g(\lambda) = \sum_{n=1}^{N} g_n^{'}(\lambda) + \lambda P_{tot}, \tag{3.29}$$

Let us assume that user k is allocated on subchannel n. Given a fixed value of λ, the object of max operation in (3.28) is a concave function of P_n. The optimality condition maximizing $g_n^{'}(\lambda)$ is given by

$$P_n = \left(\frac{B}{\log 2\lambda} - g_{k,n}\right)^{+}. \tag{3.30}$$

By searching over all K possible user assignment for subcarrier n, it is possible to evaluate $g_n^{'}(\lambda)$ and, eventually, the overall Lagrange dual function $g(\lambda)$. Then, to find $\lambda^* \geq 0$ maximizing $g(\lambda)$, an iterative approach based on the bisection method [25] requiring $\mathcal{O}(NK)$ iterations is utilized until the convergence is reached. It has been shown that the solution obtained in the dual domain becomes a global optimal solution as the number of subcarriers increases [36].

3.2.1.1 Multi-user RA with fairness

The problem of maximizing the total data rate while guaranteeing at the same time fairness among users may be formulated in several ways. Among these, Rhee and Cioffi [33] formulated the problem of assigning resources with the goal of maximizing the minimum capacity offered to each user. Mathematical formulation of this problem is

$$\max \min_{k} R_k \tag{3.31}$$

$$\sum_{k=1}^{K} a_{k,n} \leq 1 \quad n = 1, \ldots, N \tag{3.31.1}$$

$$\sum_{n=1}^{N} P_n \leq P_{tot}, \tag{3.31.2}$$

$$a_{k,n} \in \{0, 1\} \quad k = 1, \ldots, K \quad n = 1, \ldots, N \tag{3.31.3}$$

The idea is to assign more power to users with poor channel conditions so as to guarantee a data rate which can be comparable to that of users with better channel gains. Since the problem (3.31) is not convex, its optimal solution can only be found through an exhaustive search over all possible channel assignment satisfying constraints (3.31.1)-(3.31.3), leading to high computational complexity. Rhee and Cioffi [33] proposed a reduced complexity subcarrier allocation algorithm, which is based on flat transmit power. Based on this assumption, the power allocated on each subcarrier is constant and equal to $P_n = P_{tot}/N$. Let S_k be the set of allocated subchannels to user k. The subcarrier allocation algorithm is summarized by the following pseudocode (Algorithm 1):

Algorithm 1 Max-min subcarrier Allocation

$A \leftarrow \{1, \ldots, N\}$ ▷ Initialization
for $k \leftarrow 1$ to K **do**
    $R_k \leftarrow 0$, $S_k \leftarrow \emptyset$
end for
for $k \leftarrow 1$ to K **do**
    Find n satisfying $|H_{k,n}| \geq |H_{k,j}|$ for $j \in A$
    $R_k \leftarrow B \log_2 \left(1 + \frac{P_n |H_{k,n}|^2}{\sigma_w^2}\right)$
    $S_k \leftarrow S_k \cup \{n\}$; $A = A - \{n\}$
end for
while $A \neq \emptyset$ **do**
    Find k satisfying $R_k \leq R_i$ for all i, $0 \leq i \leq K$
    For the found k, find n satisfying $|H_{k,n}| \geq |H_{k,j}|$ for $j \in A$
    $R_k \leftarrow R_k + B \log_2 \left(1 + \frac{P_n |H_{k,n}|^2}{\sigma_w^2}\right)$
    $S_k \leftarrow S_k \cup \{n\}$
    $A \leftarrow A - \{n\}$
end while

When users have different requirements on data rate, a fair solution is represented by introducing proportional constraints among the users' data rates [37]. Mathematical formulation of this problem is given by

$$\max_{\{a_{k,n}\},\{P_n\}} \sum_{k=1}^{K} R_k \tag{3.32}$$

$$\sum_{k=1}^{K} a_{k,n} \leq 1 \tag{3.32.1}$$

$$\sum_{n=1}^{N} P_n \leq P_{tot}, \tag{3.32.2}$$

$$R_1 : R_2 : \ldots : R_K = \alpha_1 : \alpha_2 : \ldots : \alpha_K \tag{3.32.3}$$

$$a_{k,n} \in \{0,1\} \quad n = 1,\ldots,N \quad k = 1,\ldots,K. \tag{3.32.4}$$

In (3.32.3) $\alpha_1 : \alpha_2 : \ldots : \alpha_K$ is the set of predetermined proportional constraints where α_k is a positive real number with $\alpha_{min} = 1$ for the user with the least required proportional rate. In [37], the authors propose a suboptimal approach based on the following steps: *i*) assuming a uniform power distribution, subcarriers are assigned to each user by giving priority to the user with the least achieved proportional rate, i.e., R_k/α_k. *ii*) In the second step, the power is reallocated between the users and then among the subcarriers through water-filling to guarantee the rate proportionality among the users.

Algorithm 2 Proportional fairness subcarrier allocation

$A \leftarrow \{1, \ldots, N\}$ ▷ Initialization
for $k \leftarrow 1$ to K **do**
    $R_k \leftarrow 0$, $S_k \leftarrow \emptyset$
end for
for $k \leftarrow 1$ to K **do**
    Find n satisfying $|H_{k,n}| \geq |H_{k,j}|$ for $j \in A$
    $S_k \leftarrow S_k \cup \{n\}$
    $A \leftarrow A - \{n\}$
    $R_k \leftarrow B \log_2 \left(1 + \frac{P_n |H_{k,n}|^2}{\sigma_w^2}\right)$
end for
while $A \neq \emptyset$ **do**
    Find k satisfying $R_k/\alpha_k \leq R_i/\alpha_i$ for all i, $1 \leq i \leq K$
    Find n satisfying $|H_{k,n}| \geq |H_{k,j}|$ for all $j \in A$
    $S_k \leftarrow S_k \cup \{n\}$
    $A \leftarrow A - \{n\}$
    $R_k \leftarrow B \log_2 \left(1 + \frac{P_n |H_{k,n}|^2}{\sigma_w^2}\right)$
end while

The subcarrier allocation algorithm (Algorithm 2) tends to allocate to each user those subchannels with the highest SNRs. At each iteration, the user with the lowest proportional capacity has the possibility to pick the best subchannel. The allocation algorithm is suboptimal since it assumes equal power over all subcarriers. With the power allocation performed in the second step of the proposed algorithm, fairness is achieved among users. To find the k-th user's power P_n, Lagrange multiplier techniques [25] are used to formulate and then solve the optimization problem resulting in K nonlinear equations with K unknowns. These equations are not linear and a closed form solution does not exist, thus non-linear iterative methods such as Newton-Raphson [38] and its variants may be used.

3.2.1.2 Numerical results

In this section we evaluate the performance of the RA resource allocation algorithms that we have described in the previous subsection. We consider a single-cell OFDMA system with radius of the cell equal to 500 m. The BS has a maximum transmit power $P_{tot} = 1$ W and the number of subcarriers is $N = 192$. The data sub-carriers spacing was set to 15 kHz, giving the total transmit bandwidth of 2.88 MHz, with the centre frequency of 2 GHz. We consider a population

of static users within the cell, while the channel modeled according to the TU scenario [39].

The results presented in this section are obtained for all algorithms presented in Section 3.2.1 over 500 realizations. For the sum rate maximization with proportional rate constraints two different sets of rate constraints have been studied: the case where the rate constraints are set equal for all users (Prop rate 1), i.e. $\alpha_k = 1$ $(k = 1, \ldots, K)$, and the case where the rate constraints are set proportional to the user pathloss (Prop rate 2).

Figure 3.4 shows the average throughput for the different algorithms: the sum rate maximization algorithm achieves the highest throughput and the max-min the lowest. The results show also the flexibility of the algorithm with proportional rate constraints. As expected, when the set of rate constraints are all equal its behavior is almost identical to the max-min algorithm. On the other hand, when the system tends to favor the users nearer to the BS, the throughput approaches the sum rate results.

Figure 3.5 shows the average fairness index (according to the definition in [40]) for the various RA algorithms. In this case the max-min and the algorithm with equal rate constraints outperform all the others. The algorithm with rate constraints proportional to the pathloss even if guarantees access to all users is not very fair. This is due to the fact that in our simulation setting the difference in patlhoss can be several orders of magnitude large. Thus, users close to cell boundaries will have a much smaller throughput than users near the BS.

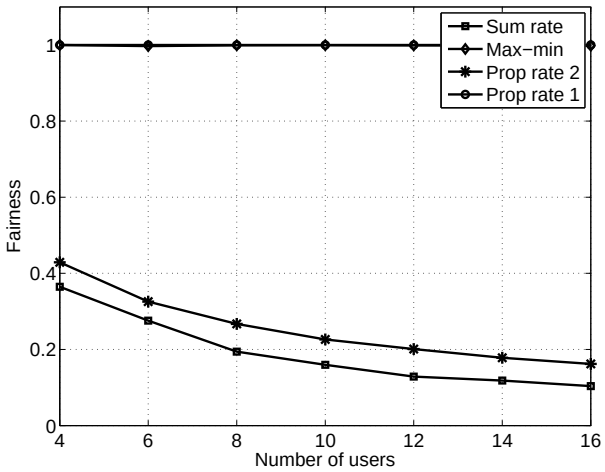

Figure 3.4: Fairness vs. number of users

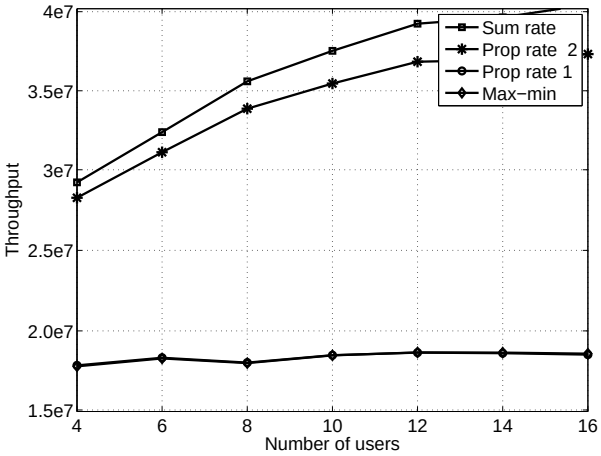

Figure 3.5: Throughput vs. number of users

3.2.2 Multi-user MA schemes

Margin-adaptive resource allocation algorithms are formulated with the goal of minimizing the overall transmitted power subject to individual user rate constraints. The mathematical formulation of this optimization problem is

$$\min_{\{a_{k,n}\},\{P_n\}} \sum_{n=1}^{N} P_n \tag{3.33}$$

$$\sum_{k=1}^{K} a_{k,n} \leq 1 \tag{3.33.1}$$

$$\sum_{n=1}^{N} r_{k,n} \geq R_{k,\min} \quad k = 1, \ldots, K \tag{3.33.2}$$

$$a_{k,n} \in \{0, 1\} \quad k = 1, \ldots, K \quad n = 1, \ldots, N \tag{3.33.3}$$

Constraints are expressed by equations (3.33.1)-(3.33.3). In (3.33.2) $R_{k,\min}$ is the minimum data rate for user k. As for RA algorithms, looking for a solution of (3.33) results into a combinatorial optimization problem, which requires an exhaustive search over all possible channel assignments. Hence, the optimal solution can be achieved by using standard optimization techniques. The Lagrangian of (3.33) is given by

$$\mathcal{L}\left\{\{P_n\}, \{r_{k,n}\}, \boldsymbol{\mu}\right\} = \sum_{n=1}^{N} P_n - \sum_{k=1}^{K} \mu_k \left(\sum_{n=1}^{N} r_{k,n} - R_{k,\min} \right). \tag{3.34}$$

The Lagrange dual function is

$$g(\boldsymbol{\mu}) = \min_{\{P_n\},\{r_{k,n}\}\in\mathcal{D}} \mathcal{L}\left(\{P_n\},\{r_{k,n}\},\boldsymbol{\mu}\right). \tag{3.35}$$

Rearranging the order of the sum terms, the minimization of (3.34) can be decomposed into N independent optimization problems as follows

$$g_n^{'}(\boldsymbol{\mu}) = \min_{\{P_n\}\in\mathcal{D}} \left\{ \sum_{k=1}^{K} P_n - \sum_{k=1}^{K} \mu_k \sum_{n=1}^{N} r_{k,n} \right\} \tag{3.36}$$

Thus, the Lagrange dual function can be rewritten as

$$g(\boldsymbol{\mu}) = \sum_{n=1}^{N} g_n^{'}(\boldsymbol{\mu}) + \sum_{k=1}^{K} R_{k,\min}, \tag{3.37}$$

Once the vector $\boldsymbol{\mu}$ is fixed, the object of the min operation is a convex function of P_n. The solution for minimization of $g_n^{'}(\boldsymbol{\mu})$ is given by

$$P_n = \left(\frac{B\mu_k}{\log 2} - \frac{1}{g_{k,n}} \right)^{+}. \tag{3.38}$$

In order to find the optimal value $\boldsymbol{\mu^*} \succeq 0$ that maximizes $g(\boldsymbol{\mu})$, the update of $\boldsymbol{\mu}$ can be efficiently performed by employing the *ellipsoid method* [25], which is a generalization of the bisection method to problems in multi-dimensional spaces. At iteration p, this algorithm finds a new ellipsoid

$$\mathcal{E}^{(p)} = \left\{ z \in \mathbb{R}^K : (z - \boldsymbol{\mu}^{(p)})^T \mathbf{A}^{(p)} (z - \boldsymbol{\mu}^{(p)}) \leq 1 \right\} \tag{3.39}$$

centered in $\boldsymbol{\mu}^{(p)}$ and with a shape defined by the positive semidefinite matrix $\mathbf{A}^{(p)}$. By construction, $\mathcal{E}^{(p)}$ contains the optimal $\boldsymbol{\mu}^*$ and its volume is smaller than that of $\mathcal{E}^{(p-1)}$, the ellipsoid found at the preceding iteration. Hence, after a certain number of iterations the ellipsoid's volume will tend to zero and its center will coincide with $\boldsymbol{\mu}^*$. Ellipsoid method involves $\mathcal{O}(n^2)$ iterations, with n denoting the number of variables [25] and utilizes as a subgradient the following vectors:

$$d_k = R_{k,\min} - \sum_{n=1}^{N} r_{k,n}^* \quad k = 1, \ldots, K \tag{3.40}$$

where $r_{k,n}^*$ and P_n^* are the optimal solution for problem (3.35) for a fixed $\boldsymbol{\mu}$. The overall optimization needs $\mathcal{O}(K^2)$ iterations of optimization problem with complexity of $\mathcal{O}(NK)$. Thus, $\mathcal{O}(NK^3)$ executions are required to find the optimal solution of problem (3.33).

Another way to solve problem (3.33) is proposed by Wong *et al.* in [26]. Here the authors find a suboptimal solution which is based on the relaxation of the integrality condition on the

allocation variable $a_{k,n}$. In this case, the latter can also be interpreted as the time-sharing factor for user k on subcarrier n. The allocation problem is reformulated as

$$\min_{\{a_{k,n}\},\{P_{k,n}\}} \sum_{k=1}^{K}\sum_{n=1}^{N} a_{k,n}P_{k,n} \tag{3.41}$$

$$\sum_{k=1}^{K} a_{k,n} \leq 1 \qquad n = 1,\ldots,N \tag{3.41.1}$$

$$\sum_{n=1}^{N} \rho_{k,n} \geq R_{k,\min} \quad k = 1,\ldots,K \tag{3.41.2}$$

$$a_{k,n} \in [0,1] \quad n = 1,\ldots,N \quad k = 1,\ldots,K. \tag{3.41.3}$$

where $P_{k,n}$ is the power transmitted by the BS to the user k on subcarrier n, whereas in equation (3.41.1) $\rho_{k,n} = a_{k,n}r_{k,n}$ is a new rate variable and $a_{k,n}$ can take all values within the interval $[0,1]$. With this new parameter, the optimization problem can be reformulated as a convex minimization problem over a convex set. Moreover, using Shannon formula, $P_{k,n}$ is function of $\rho_{k,n}$ and $a_{k,n}$ as

$$P_{k,n} = (2^{\frac{\rho_{k,n}}{a_{k,n}}} - 1)\frac{1}{g_{k,n}} \tag{3.42}$$

Thus, the Lagrangian of problem (3.41) is

$$\begin{aligned}\mathcal{L}(\{\rho_{k,n}\},\{a_{k,n}\},\boldsymbol{\lambda},\boldsymbol{\mu}) = &\sum_{k=1}^{K}\sum_{n=1}^{N}\left(2^{\frac{\rho_{k,n}}{a_{k,n}}} - 1\right)\frac{a_{k,n}}{g_{k,n}} \\ &- \sum_{k=1}^{K}\lambda_k\left(\sum_{n=1}^{N}\rho_{k,n} - R_{k,\min}\right) - \sum_{n=1}^{N}\mu_k\left(\sum_{k=1}^{K}\rho_{k,n} - 1\right)\end{aligned} \tag{3.43}$$

After differentiating $\mathcal{L}(\{\rho_{k,n}\},\{a_{k,n}\},\boldsymbol{\lambda},\boldsymbol{\mu})$ with respect to $\rho_{k,n}$ and $a_{k,n}$, we obtain the necessary conditions for the optimal solutions, $\rho^*_{k,n}$ and $a^*_{k,n}$. Given a fixed set of Lagrange multipliers λ_k, $k = 1,\ldots,K$, we have

$$\frac{\rho_{k,n}}{a_{k,n}} = \begin{cases} 0 & \lambda_k \leq \log 2/g_{k,n} \\ \log_2(\lambda_k g_{k,n}/\log 2) & \lambda_k \leq \log 2/g_{k,n} \end{cases} \tag{3.44}$$

and

$$a^*_{k',n} = \begin{cases} 1 & k = \arg\min\{\tau(\lambda_k, g_{k,n})\} \\ 0 & \text{otherwise} \end{cases} \tag{3.45}$$

where

$$\tau(\lambda_k, g_{k,n}) = \frac{\lambda_k g_{k,n} - \log 2}{\log 2 \cdot g_{k,n}} - \lambda_k \log_2\left(\frac{\lambda_k g_{k,n}}{\log 2}\right) \tag{3.46}$$

In order to determine the set of λ_k so as to satisfy the users' rate constraints, Wong *et al.* follow an iterative approach that is performed for each user in the system. This procedure repeats for

all users until the data rate constraint for all users are satisfied. The obtained set of Lagrange multipliers determines the optimal sharing factor of all the subcarriers for all users. However, since each subcarrier must be assigned to only one user, a suboptimal two-step procedure in which subcarrier and power allocation are performed separately, is derived. The subcarriers are exclusively allocated to the user having the largest time-sharing factor on that subcarrier. After that, power allocation is performed independently for each user over the assigned subchannel.

3.2.2.1 LP approach

Another way to solve resource allocation problems is to formulate them as a LP problem, where both the objective function and the constraints are linear [41]. Let us assume that all user requests are expressed as an integer multiple of a certain fixed rate $R_0 = B\eta_0$, i.e. the rate requested by user k is $R_{k,\text{target}} = n_k R_0$, with n_k an integer number. We also assume that the transmission rate of a subcarrier is $Bf\eta_0$, where $f \in \mathcal{F}$ is the adopted transmission format and $\mathcal{F} = \{1, \ldots, F\}$ is the set of available format. Therefore, according to the Shannon formula, the power required to support the rate fR_0 on subcarrier n is given by

$$P_{k,n,f} = \left(2^{f\eta_0} - 1\right) \frac{1}{g_{k,n}} \tag{3.47}$$

Let us define $a_{k,n,f}$ as the integer variable that is set to 1 if the n-th subchannel is assigned to the k-th user at transmission format f, and 0 otherwise. The resource allocation problem can be formulated as

$$\min_{\{a_{k,n,f}\},\{P_{k,n,f}\}} \sum_{k=1}^{K} \sum_{n=1}^{N} \sum_{f=1}^{F} a_{k,n,f} P_{k,n,f} \tag{3.48}$$

$$\sum_{k=1}^{K} \sum_{f=1}^{F} a_{k,n,f} \leq 1 \qquad n = 1, \ldots, N \tag{3.48.1}$$

$$\sum_{n=1}^{N} \sum_{f=1}^{F} f\eta_0 a_{k,n,f} = R_{k,\text{target}} \quad k = 1, \ldots, K \tag{3.48.2}$$

$$a_{k,n,f} \in \{0, 1\} \quad n = 1, \ldots, N \quad k = 1, \ldots, K. \tag{3.48.3}$$

The problem (3.48) belongs to the class of LIP. The solution of such problems can be achieved through an exhaustive search over all possible assignment of subcarrier, power and transmission format. However, in many cases, the computational complexity could be too high. A common solution is to follow the approach of [26] by letting $a_{k,n,f}$ vary continuously in the interval $[0, 1]$. By doing so, the problem (3.48) becomes a standard LP problem and can be solved by using

any of the algorithms used for solving LP problems (e.g. the IPM [42]), which are able to find the solution of large scale problems with a polynomial time.

In order to reduce the overall complexity, we modify the problem by choosing a single transmission format for all users independently of their channel gains. With this choice, problem (3.48) can be restated as

$$\min_{\{a_{k,n}\},\{P_{k,n}\}} \sum_{k=1}^{K}\sum_{n=1}^{N} a_{k,n}P_{k,n} \tag{3.49}$$

$$\sum_{k=1}^{K} a_{k,n} \leq 1 \qquad n = 1,\ldots,N \tag{3.49.1}$$

$$\sum_{n=1}^{N} a_{k,n} = n_k \quad k = 1,\ldots,K \tag{3.49.2}$$

$$a_{k,n} \in \{0,1\} \quad n = 1,\ldots,N \quad k = 1,\ldots,K. \tag{3.49.3}$$

Again, relaxing the integrality condition on $a_{k,n}$ leads the problem (3.49) to become a LP problem. It can be shown that the exact solution of this LP problem is equal to the solution of LIP problem (3.49) [43]. Moreover, the relaxed LP problem has the characteristic that can be modeled as a *network flow* problem [44] and it can therefore be solved with very fast algorithms [45] . Among these, the NSM os the most efficient solver for min-cost-max-flow network problem and runs in polynomial time [46].

The single-format solution allows a great simplification of the solution of the problem (3.48) at the cost of a modest worsening of system performance. This loss is partially compensated by the multiuser diversity, which is well exploited by the dynamical assignment of resources [47].

3.2.2.2 Numerical results

We analyze the performance of the MA resource allocation algorithms. In the following we refer to as Ideal bound the algorithm obtained via Lagrangian dual decomposition, whereas WCLM corresponds to the algorithm proposed by Wong *et al.* [26]. We consider a single-cell OFDMA system with radius of the cell equal to 500 m. The BS transmit over a downlink bandwidth $W = 5$ MHz thus the sampling time is $T_c = 200$ ns and the number of subcarriers is $N = 16$, with the centre frequency of 2 GHz. The pathloss of the channel is proportional to the distance between the BS and the MS and the pathloss exponent is 4. The channel is modeled as static and frequency-selective Rayleigh fading with an exponentially decaying power delay profile. The channel delay spread is $\sigma_\tau = 1\mu s$ and the coherence bandwidth is assumed to be larger

than the bandwidth spanned by each subchannel. Each user has the same rate requirement $R_{k,\text{target}} = R_0,\ k = 1, \ldots, K$. Furthermore, we consider a population of static users within the cell. The results shown in the following have been obtained by averaging on 100 channel realizations.

Figure 3.6 shows the mean transmitted power P_m expressed in W as a function of the of the spectral efficiency per subcarrier η expressed in bit/s/Hz. The number of users is $K = 4$. We see that as the cell load increases, the power consumption increases as well in order to satisfy the individual rate requirements for each user. The LP approach exhibits only a limited increasing of power consumption, compared to the other algorithms. However, this loss is compensated by the reduction of computational complexity provided by use of LP methods. Figure 3.7 illustrates the mean transmitted power P_m as a function of the number of users K.

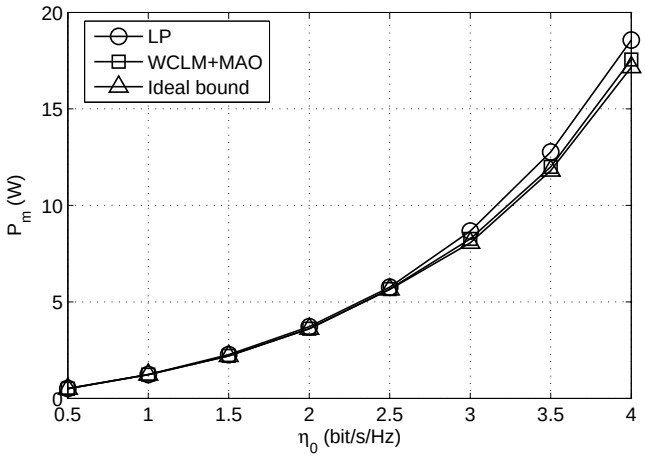

Figure 3.6: Mean transmitted power P_m vs. spectral efficiency η

We have set a target spectral efficiency per subcarrier $\eta = 2$ bit/s/Hz. As the number of users increases, the transmitted power reduces due to the exploiting of multi-user diversity of the system. Indeed, with a larger number of MSs in the cell, the BS has more degree of freedom to allocate power and subcarrier, thus the probability that a user is given a channel with poor quality decreases. Moreover, the gap between LP and the optimal bound as well as the WCLM algorithm is always more negligible as K increases.

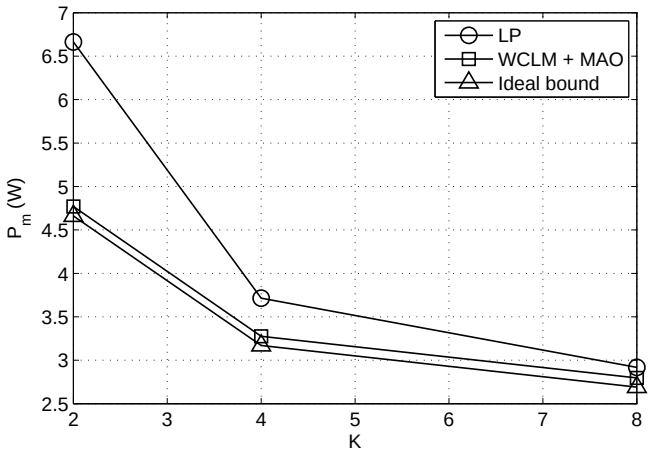

Figure 3.7: Mean transmitted power P_m vs. number of users K

Chapter 4

Resource allocation in multi-cell OFDMA systems

In multicellular OFDMA systems, one of the most limiting factor on wireless capacity is represented by high level of MAI, caused by users that share the same available frequency spectrum. To reduce the impact of MAI, most conventional cellular systems employ orthogonal signaling within a cell and a frequency reuse distance larger than one. This strategy trades spectral efficiency for robustness to interference and requires either accurate frequency planning or an explicit cooperation among cells. Systems based on CDMA do indeed allow a reuse factor equal to one [48] at the cost of a penalty in terms of cell capacity when a large number of users is active in neighboring cells. Only recently did standardization activities [49] focus on cellular architectures based on OFDMA with a full reuse of the frequency spectrum.

The problem of resource allocation in multi-cellular reuse-one OFDMA systems has been only recently addressed in the literature. The general problem of maximizing a weighted sum of the users' rates in a mutual interference scenario is shown to be NP-hard in [50]; thus, no low-complexity algorithms exist that return the optimal solution. Efficient suboptimal solutions for resource allocation in multicell OFDMA systems are discussed in [51], where coordinated strategies are shown to outperform uncoordinated ones. A hybrid centralized-distributed scheme, where resource allocation decisions are jointly taken by the RNC and the BSs, is proposed in [52]. The authors in [53] introduce a distributed scheme, in which the allocation is performed independently in each cell with the goal of minimizing overall power consumption given the user rate requirements. A centralized iterative algorithm, based on a central controller allocating

resources to clusters of cells, is introduced in [54]. The proposed heuristic is reasonably fast and achieves a good efficiency. While centralized strategies can indeed achieve a higher resource utilization, distributed solutions are simpler to implement [55].

Reuse-one systems tend to penalize users on the cell border, which receive a relatively weak signal from their BS and are subject to a higher interference coming from other cells. Thus, FFR systems [56] statically partition the cell into an inner, reuse-one area, and an outer area, which is divided into sectors, so that different frequencies are allocated to adjacent sectors.

In this Chapter we consider the downlink of an OFDMA multi-cellular system where each BS is assumed to have perfect knowledge of the CSI of the users in its cell only. Radio resources are allocated with the goal of minimizing the transmitted power subject to individual user rate constraints. As in [47] the latter can be dynamically varied over time so as to achieve a higher energy efficiency, thus leading to some level of unfairness in the short term. Fairness in the long term is enforced by a *packet scheduler*, which selects on each time interval a subset of scheduled users, as well as their maximum data rate. Channel assignment is formulated as a linear programming problem, which can be solved as a minimum cost network flow problem with limited complexity. Allocation is performed independently in each cell, resulting in an *iterative scheme*: whenever a cell modifies its allocation it perturbs the interference perceived in neighboring cells, which accordingly need to change their own allocation patterns. After some iterations the system may converge to a *steady state* for a given traffic pattern and channel realization, and the allocations become stable. To guarantee convergence, we define a fully distributed and adaptive *load control* algorithm.

A distributed resource allocation strategy such as the one we propose could, in principle, lead to a significant performance degradation in comparison with centralized strategies. We show that this is not the case, by comparing the performance of our distributed resource allocation algorithm with an approximation of a lower bound of an equivalent centralized optimization for resource allocation. The performance loss appears to be negligible, provided that there are more than a few users in the system. The reason is mainly the efficient exploitation of user and channel diversity allowed by our proposed resource allocation architecture (scheduler+allocator+load control).

4.1 System model and problem formulation

We consider a cellular system composed of N_{cells} cells that share the whole available spectrum, i.e., the frequency reuse is equal to one. In each cell i are a BS and $K^{(i)}$ MSs. The BS schedules the users and assigns them a subset of the radio channels. The modulation technique is multi-carrier and the multiple access scheme is OFDMA, so that within each cell the system bandwidth is partitioned into N orthogonal subcarriers and each user is assigned a different subset of subcarriers. To perform resource allocation, the BS needs to know channel gains and interference levels on all subcarriers for all users in the cell. As shown in Chapter 3, to reduce allocation complexity, we group sets of adjacent subcarriers into sub-channels of bandwidth B.

The time axis is organized into *frames*: as illustrated in Figure 4.1, each frame is composed of an *allocation phase* and a *transmission phase*. In the allocation phase, that has a duration of T_a seconds and is is made up of N_a slots of duration T_s seconds, each cell converges to a stable allocation. The value of N_a is a system-wide constant. During the T_t seconds of the transmission phase, the users transmit their data on the channels that have been assigned to them during the allocation phase. Scheduling is performed on a frame basis.

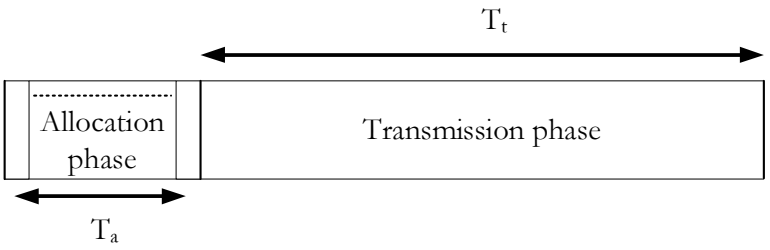

Figure 4.1: Frame structure

Our goal is to minimize the overall power transmitted in each cell subject to a rate constraint per user: each BS i needs to assign to each served user k a set of orthogonal channels $S_k^{(i)}$ and distribute power in such a way that all rate requirements $R_k^{(i)}$ $(k = 1, \ldots, K^{(i)})$ are met. By indicating with $P_n^{(i)}$ the power transmitted on sub-channel n in cell i, and using the Shannon capacity as a measure of the transmitted rate, the allocation problem can be formulated as a

set of N_{cells} optimization problems. In cell i the problem is stated as:

$$\min \sum_{k=1}^{K^{(i)}} \sum_{n \in S_k^{(i)}} P_n^{(i)} \tag{4.1}$$

$$\sum_{n \in S_k^{(i)}} B \log_2(1 + \gamma_{k,n}^{(i)}) \geq R_k^{(i)} \qquad k = 1, \ldots, K^{(i)} \tag{4.1.1}$$

$$S_j^{(i)} \cap S_k^{(i)} = 0 \qquad \forall j \neq k. \tag{4.1.2}$$

User rate requirements are set in (4.1.1); the received SINR for user k on sub-channel n in cell i, $\gamma_{k,n}^{(i)}$, is computed as

$$\gamma_{k,n}^{(i)} = \frac{P_n^{(i)} \left| H_{k,n}^{(i)} \right|^2}{\sigma_w^2 + I_{k,n}^{(i)}} \tag{4.2}$$

where $H_{k,n}^{(i)}$ is the channel gain between user k and the BS i on the n-th link, and σ_w^2 is the power of the zero-mean thermal noise. In the following we are considering a population of slow-moving users so that the propagation channel has a long coherence time. The MAI $I_{k,n}^{(i)}$, affecting user k in the cell i on the n-th channel, is given by

$$I_{k,n}^{(i)} = \sum_{\substack{j=1 \\ j \neq i}}^{N_{cells}} P_n^{(j)} \left| H_{k,n}^{(j)} \right|^2 \tag{4.3}$$

Constraints (4.1.2) impose that resource assignment is orthogonal within a cell.

The solution of the N_{cells} optimization problems is not trivial. We have showed in Chapter 3 that even in the single-cell case, i.e. $N_{cells} = 1$, problem (4.1) is not convex and can be solved either by relaxing the exclusivity constraints on sub-channel allocation [26] or, if the duality gap is zero, by finding the solution of the Lagrangian dual problem [35]. In both cases, finding the allocation is a computationally demanding strategy. Moreover, in our case, due to the interference, the allocation in one cell perturbs the allocations in all other cells and the effect of MAI is such that, given the user rate constraints, an allocation solution that meets all the requirement may not exist.

Assuming lack of any centralized control, the presence of strong interference caused by full frequency reuse suggests the implementation of iterative strategies [57]. In the following we propose a distributed algorithm where each cell independently solves problem (4.1) on the base of the interference it measures and iteratively updates its allocation with the goal of satisfying all its rate requirements. The problem of convergence is addressed by progressively reducing the traffic load of the cells which can not find a stable allocation; fairness is achieved by

implementing a *layered architecture* [47], where radio resource management and traffic policies are dealt with by separate and interacting algorithms [1].

4.2 Distributed layered allocation architecture

The DLA is composed of a PS and a RRA, both located at the BS. The scheduler works on a frame-by-frame basis: at the beginning of a new frame it dictates to the allocator the new set of rate requirements for all active users. Due to the MAI, the RRA may require a number of iterations to find an allocation matching the scheduler requirements. The practical choice of the number of iterations in the *allocation phase* is a crucial issue since it has an impact on throughput, overhead, and power.

4.2.1 Single-cell radio resource allocation based on linear programming

Following the same approach as in Section 3.2.2.1, we reformulate channel assignment (4.1) as a LP problem and adopt a single transmission format (meaning error correction code and modulation) for all users. Such a choice allows a great simplification of the allocation problem and, as shown in Section 4.3, has only a minor impact on allocation performance since most of the channel diversity is exploited by dynamically assigning channels to users [43]. Thus, the request of a given rate $R_k^{(i)}$ directly translates into the request of a certain number of sub-channels $n_k^{(i)}$ $(k = 1, \ldots, K^{(i)})$.

A given transmission format corresponds to a certain spectral efficiency $\eta = \log_2 [1 + \gamma(\eta)]$, where $\gamma(\eta)$, the target SINR to achieve the spectral efficiency η, is $\gamma(\eta) = 2^\eta - 1$. Assuming perfect channel estimation, user k allocated in cell i on sub-channel n transmits with rate $R = B\eta$ if its measured SINR is $\gamma_{k,n}^{(i)} > \gamma(\eta)$ and with rate $R = 0$ otherwise. Thus, the power necessary for user k in cell i to transmit on sub-channel n with spectral efficiency η is computed from (4.2) as

$$P_{k,n}^{(i)}(\eta) = \gamma(\eta) \frac{BN_0 + I_{k,n}^{(i)}}{\left| H_{k,n}^{(i)} \right|^2}. \tag{4.4}$$

[1] A similar decomposition of the resource allocator is typical of OFDMA systems; e.g., in LTE [3] a time-domain scheduler selects the users to be scheduled in the next TTI, while a frequency-domain scheduler allocates resources to the selected users.

Note that $I_{k,n}^{(i)}$ is the MAI as measured in the previous iteration of the allocation phase. Under these hypotheses, once the transmission format is set, each sub-channel has a certain fixed cost for each user and the allocation problem consists in finding the channel assignment that minimizes the sum of the costs in each cell. Thus, introducing the binary allocation variable $a_{k,n}^{(i)}$, which is 1 if sub-channel n is assigned to user k in cell i and 0 otherwise, given a certain η, the resource allocation problem in cell i is formulated as follows:

$$\min_{\mathbf{a}^{(i)}} \sum_{n=1}^{N} \sum_{k=1}^{K^{(i)}} P_{k,n}^{(i)}(\eta) a_{k,n}^{(i)} \tag{4.5}$$

$$\sum_{n=1}^{N} a_{k,n}^{(i)} \leq n^{(i)}(k) \qquad k = 1, \ldots, K^{(i)} \tag{4.5.1}$$

$$\sum_{k=1}^{K^{(i)}} a_{k,n}^{(i)} \leq 1 \qquad n = 1, \ldots, N \tag{4.5.2}$$

$$\sum_{k=1}^{K^{(i)}} \sum_{n=1}^{N} b_{k,n}^{(i)} = C_{req}^{(i)} \tag{4.5.3}$$

$$a_{k,n}^{(i)} \in \{0, 1\} \quad n = 1, \ldots, N \quad k = 1, \ldots, K^{(i)} \tag{4.5.4}$$

After allocation, the total rate assigned to user k in cell i is

$$R_k^{(i)} = \sum_{n=1}^{N} a_{k,n}^{(i)} B\eta. \tag{4.6}$$

Constraints (4.5.1) formulate the rate requirements per user. Following a different strategy from that in (4.1), where each user is required to achieve *at least* a certain rate, this set of constraints imposes that each user gets *at most* a certain number of resources. The PS policy is thus enforced: with the goal of guaranteeing long-term fairness, $n_k^{(i)}$ will be large for those users that have received a small share of resources and small or even zero for those users that have already received a large amount of radio resources. Without the limits imposed by constraints (4.5.1), the RRA would naturally assign all the resources to the users with the best channels only. Moreover, by setting a *soft* constraint, such as the maximal amount of resources, the PS gives the RRA enough freedom to exploit the diversity of the system. Constraints (4.5.2) implement orthogonal access within the cell, i.e., one sub-channel can be assigned to only one user per cell; they mirror constraint (4.1.2). Constraint (4.5.3) indicates the total amount of resources to assign in the cell and prevents the allocator from choosing the solution where all variable allocations are set to 0. Initially, it is $C_{req}^{(i)} = N$ in every cell; then if, due to the MAI, the system is not able to converge to a steady allocation, $C_{req}^{(i)}$ is progressively reduced by the

load control algorithm. As to (4.5.4), the allocation variable is integer and can only assume the values 0 and 1; hence, we are dealing with an LIP problem. LIP problems are in many practical situations NP-hard. Nevertheless, we prove the following Theorem resulting from the *total unimodularity* property of the constraint matrix.

Theorem 1. *The LP problem obtained from (4.5) by removing the integrality condition (4.5.4) has an* integral *optimal solution.*

Proof. By replacing the equality constraint (4.5.3) with two inequalities the constraints of problem (4.5) can be written in the form:

$$\mathbf{Bx} \leq \mathbf{c} \tag{4.7}$$

where $\mathbf{x} = [a_1(1) \dots a_K(1) \ \dots \ a_1(N) \dots a_K(N)]'$, $\mathbf{c} = [r(1) \dots r(K) \ 1 \dots 1 \ C_{req}^{(i)} \ -C_{req}^{(i)}]'$. The constraint matrix $\mathbf{B}$ has $N + K + 2$ rows and $K \cdot N$ columns, and can be written in the form $[\mathbf{B_1'} \ \mathbf{B_2'} \ \mathbf{1'_{N\times K}} \ -\mathbf{1'_{N\times K}}]'$, where $\mathbf{B_1}$ is the matrix corresponding to the constraints (4.5.1) and $\mathbf{B_2}$ is the matrix corresponding to the constraints (4.5.2); the two inequality constraints corresponding to (4.5.3) are represented by a vector of ones and a vector of minus ones, respectively. We show that $\mathbf{B}$ is a *totally unimodular* matrix, i.e. every square submatrix of $\mathbf{B}$ has determinant equal to 0, 1 or -1. We proceed by proving that for each collection R of rows in $\mathbf{B}$ we can always find a partition $R = R_1 \dot{\cup} R_2$, such that:

$$\sum_{i \in R_1} b_{ij} - \sum_{i \in R_2} b_{ij} \in \{-1, 0, 1\} \qquad \forall j \in \{1, \dots, K \cdot N\}\,.$$

These are necessary and sufficient conditions for the unimodularity of matrix $\mathbf{B}$ (see [58]). We distinguish four cases:

1. if R does not contain the last two rows of $\mathbf{B}$, then we can put the rows of R belonging to $\mathbf{B_1}$ in R_1 and those belonging to $\mathbf{B_2}$ in R_2;

2. if R contains the $N + K + 1$-th row of $\mathbf{B}$, i.e., the vector $\mathbf{1_{N\times K}}$, but *not* the last row, we put the rows belonging to $\mathbf{B_1}$, $\mathbf{B_2}$ in R_1 and let $R_2 = \{N + K + 1\}$;

3. if R contains the last row of $\mathbf{B}$, i.e., the vector $-\mathbf{1_{N\times K}}$, but *not* the penultimate row, we put all the rows in R_1 and let $R_2 = \emptyset$;

4. if R contains the last two rows of $\mathbf{B}$, we put in R_1 the rows belonging to $\mathbf{B_1}$, and in R_2 all the remaining rows.

Since **B** is totally unimodular matrix [46], and the constraint bounds of the problem are integer, the polyhedron defined by the constraint set has integer vertices, and the LP problem has an integral optimal solution [42]. □

Hence, LP methods can be used to solve (4.5) with very limited complexity [45], [46].

The iterative implementation of resource allocation brings up two major problems: the allocation complexity and the overhead exchanged to update the resource costs. While the efficient formulation of the problem is intended to address the former issue, the latter needs to be discussed in detail. For each iteration of the allocation phase the RRA needs updated knowledge of the interference power on each sub-channel for each user in the cell. Thus, with the objective to reduce the amount of signaling overhead, we assume that adjacent BSs are interconnected by high-capacity links, such as the X2 interface [59][60], supported in LTE. Under this hypothesis, the allocation phase conforms to the following iterative procedure

1. *Initialization.* Each user estimates the gains of the channels with its serving BS and all neighboring BSs and signals this information to its BS. On the base of this information, each BS *virtually* performs allocation solving the LP problem (4.5) and signals *through the high-capacity links* to the other BSs the power values that intends to use on each channel;

2. *Iteration.* At the beginning of a new iteration, each BS is able to compute for all its served users the interference and hence the power cost for each subcarrier. Having updated the costs, the BSs newly solve the allocation problem (4.5) and signal to all neighboring BSs the new power vector until convergence.

The only information that is actually exchanged in the air between users and BS during the allocation phase is the vector of estimated channel gains at initialization. In practical scenarios, the overhead can be further reduced by signaling only the gains relative to the most interfering BS. All other information is exchanged among BSs through the high speed link. In this case the overhead amounts to signaling the vector of power values that a BS intends to transmit for each iteration. Although the BSs do exchange information, the allocation procedure remains *distributed* in the sense that each BS takes its allocation decisions autonomously from the others. At the end of the allocation phase each BS broadcasts the resource assignment for each user and the actual data transmission takes place. Moreover, since most of the signaling is exchanged through the connections between the BSs, the allocation phase of a new frame can

be performed simultaneously with the end of the transmission phase of the preceding frame, so that at the beginning of a new frame the allocation is already set.

4.2.2 Load control

Depending on the traffic configuration, the distributed RRA scheme may not achieve convergence. During the allocation phase, cells are free to change their allocation from iteration to iteration and once a cell modifies its allocation, it also changes the interference pattern it generates in the neighboring cells. Since different cells react simultaneously, this may lead to a new unstable configuration. For this reason we implement a LC mechanism that reduces the cell load so as to ease convergence to a stable allocation.

To avoid the risk of starving cell-edge users, rather than acting on single users, the LC algorithm controls the load of an entire cell. By reducing the value of $C_{req}^{(i)}$ in constraint (4.5.3), the LC algorithm reduces the total number of radio channels that the RRA can allocate in cell i and, as a consequence, also the overall level of interference that cell i generates in the system. Because of the distributed nature of the proposed architecture, whenever a cell reduces its load, neighboring cells respond by updating their allocation. By progressively reducing the load of the cells, the iterative DLA is able to self adapt to the current level of traffic, hence interference and eventually find a stable allocation.

Rather than performing a *hard* partitioning of the radio resources as in conventional systems, the layered combination of RRA, LC and PS implements a *soft* partitioning. Where possible, all cells in the system transmit over the same bandwidth, otherwise they use only a subset of the available radio channels. The cardinality of the subset is set by the LC algorithm while the actual channel assignment is performed by the RRA. Each cell performs LC autonomously.

The pseudocodes of Algorithms 3-4 shows the details of a heuristic implementation of the combination of LP-based RRA and LC (linear programming resource allocator, LPRA). In cell i at the beginning of each frame, the PS passes to the allocator the rate requirements for all users together with the load $C_{req}^{(i)}$ determined in the previous frame. The allocation phase lasts N_a iterations, and is divided into two subphases:

1. For the first N_s iterations, each cell performs sub-channel and power allocation (lines 7-22). On each iteration the BS checks whether some allocations are still oscillating. If this is the case, it reduces its load[1] with a probability ϵ proportional to the amount of

[1]Except for the first N_{ini} iterations of the frame, in which the system is left to evolve freely without any load reduction.

instability the cell brings into the system[1] (line 15). Once a cell has reduced its load, no other reduction is performed *in that cell* for the next N_p iterations (resource allocation and transmission power continue to be updated). This allows the system to evolve and possibly reach a stable configuration, while preventing an excessive load reduction in a single cell.

2. Those sub-channels on which the target SINR value has not yet been reached after N_s iterations are switched off; channel assignment is frozen, while power control is allowed to evolve for the next N_{pow} slots (lines 23-33).

Since those sub-channels on which the target SINR has not been achieved are switched off, the iterative power control in lines 34-38 always starts from a feasible power vector, and is therefore guaranteed to converge to a fixed point [61]. The following Claim thus holds:

Claim 1. *At the end of the allocation phase, a stable and feasible power assignment is achieved.*

Proof. After the first N_s iterations (lines 7-22) the allocation of sub-channels to the users will no longer be modified for the rest of the frame. At this point the system may already have reached a state in which the target SINR is achieved on all sub-channels; thus, the current power allocation would already be feasible. If this is not the case, those sub-channels on which the target SINR has not been achieved are switched off (lines 23-33); this also results in a feasible power allocation for the remaining sub-channels. Hence, the iterative power control in lines 34-38 always starts from a feasible power vector, and is guaranteed to converge to a fixed point [61]. □

To speed up convergence, the initial value of $C_{req}^{(i)}$, namely $C_{req,0}^{(i)}$ is by default the value obtained at the end of the preceding signaling phase. Our algorithm performs a reduction of the cell load in order to make it easier to achieve a feasible allocation of resources in adjacent cells using the same frequency band. Since the load reduction algorithm is distributed and has a random component, the cell load may be reduced more than would be necessary; for this reason, and in order to track varying channel and traffic conditions, a cell should be able to try increasing its load. Thus, in line 4 cell i increases its initial load value $C_{req,0}^{(i)}$, but only if no sub-channels had to be switched off in the previous frame.

[1] The value of ϵ is the ratio between the sum of the transmission powers on the sub-channels whose allocation is still oscillating, and the total transmission power in the cell.

Algorithm 3 Resource allocation and load control, Part I

for $cell \leftarrow 1$ to N_{cells} **do**
  $prIts(cell) \leftarrow 0$
  if $failedSIR(cell) = 0$ **then**
    $C_{req}(cell) \leftarrow \min\left(C_{req}(cell) + 1, C_{req}^{MAX}\right)$
  end if
end for
for $iteration \leftarrow 1$ to N_s **do**
  for $cell \leftarrow 1$ to N_{cells} **do**
    Perform resource and power allocation for $(cell, iteration)$
    if allocation still changing after N_{ini} iterations **and** $prIts(cell) = 0$ **then**
      $oscSC \leftarrow$ Set of sub-channels which have still not achieved a stable allocation
      $\epsilon \leftarrow \left(\sum_{n \in oscSC} P_n^{(cell)}\right) / \left(\sum_{n=1}^{N} P_n^{(cell)}\right)$
      $h \leftarrow$ **rnd**
      if $h < \epsilon$ **then**
        $C_{req}(cell) \leftarrow C_{req}(cell) - 1$ ▷ Reduce load with probability ϵ
        $prIts(cell) \leftarrow N_p$ ▷ Load reduction will not be performed for the next N_p iterations
      end if
    else
      $prIts(cell) \leftarrow \max\left(prIts(cell) - 1, 0\right)$
    end if
  end for
end for

Algorithm 4 Resource allocation and load control, Part II

for $cell \leftarrow 1$ to N_{cells} **do** ▷ Switch off the sub-channels below the target SINR
  $failedSIR(cell) \leftarrow 0$
  for $subCar \leftarrow 1$ to N **do**
    Keep the allocation of the last iteration for $(cell, subCar)$
    if checkTargetSIR$(cell, subCar) == 0$ **then**
      Switch off the sub-channel $subCar$ in cell $cell$
      $C_{req}(cell) \leftarrow C_{req}(cell) - 1$
      $failedSIR(cell) \leftarrow 1$
    end if
  end for
end for
for $iteration \leftarrow N_s + 1$ to $N_s + N_{pow}$ **do** ▷ Adjust final TX powers
  for $cell \leftarrow 1$ to N_{cells} **do**
    Adapt TX power, without modifying the allocation
  end for
end for

4.2.3 Credit-based packet scheduling

The goal of the PS is to enforce long-term fairness so that the users in the cell achieve throughputs proportional to *weights* assigned to them. We adopt a credits-based packet scheduling algorithm, loosely based on CBFQ [62], but expressly designed for the OFDMA radio interface. The PS schedules data blocks by taking account of the credits and weights of each user: the credits increase when the user is not scheduled to transmit, while they decrease when a transmission resource (sub-channel) is assigned to it. A novel feature of the scheduler is that the allocator is allowed some degree of freedom in deciding how many transmission resources should be assigned to each user; in fact, the scheduler only determines a maximum rate constraint per user. This leads to a more efficient allocation, while guaranteeing long-term fairness. In order to determine the maximum rate constraints for each user, the scheduler generates a list of data blocks (packets in the following); one such packet carries $L = BT_t\eta$ bits, corresponding to the amount of data that can be sent on a sub-channel during the transmission phase (B is the channel bandwidth, T_t the duration of the transmission, η the spectral efficiency).

In the following we briefly describe the scheduler and its properties; a detailed treatment can be found in [63]. At the beginning of the transmission phase, the PS selects a list of up

to C_{max} packets to be sent to the RRA. The RRA allocates up to C_{req} resource units, where $C_{req} \leq C_{max}$; the rate constraint $n^{(i)}(k)$ (4.5.1) for user k is given by the number of packets belonging to user k in the list generated by the scheduler. Note that this list is longer than what can actually be transmitted in a frame. The number of packets in the list which belong to user i constitute the upper bound on the amount of resources that can be allocated to that user. The excess packets $C_{max} - C_{req}$ represent the amount of "freedom" given to the RRA to take advantage of user diversity by exploiting some short-term unfairness. Thus the allocator selects a subset of users for transmission, and for each flow the sub-channel(s) to be assigned to it. Finally, at the end of the frame the PS receives from the RRA the list of the actually allocated packets; it is thus able to keep track of the state of each user and to correctly update the internal state associated to it. The same rate constraints are imposed when allocating sub-channels to the users during the slots which make up the allocation phase; the scheduler, however, need not take this into account, since no actual data transfer takes place.

Our goal is to fairly allocate the transmission resources to the users according to their weight. An important property of our scheduling algorithm is that this fairness goal is attained *independently* of the algorithm used by the allocator, i.e., of the policy according to which C_{req} out of C_{max} packets are selected for transmission. In fact, it can be shown [47] that the (weighted) discrepancy between the transmission rates of any two users can be made arbitrarily small by choosing a sufficiently long time interval, whatever the values of C_{max} and C_{req} used in each frame. This remains true even if they are varied over time, provided the difference $C_{max}(\tau) - C_{req}(\tau)$ can be uniformly bounded over time.

4.3 A centralized allocator for multicellular multi-carrier systems

To reduce the complexity of the allocator we use a single transmission format per user. To support this choice, we develop a benchmark model, by defining a CRA that *jointly* solves the N_{cells} allocation problems (4.1) and can be formulated as follows

$$\min_{\mathbf{P},\mathbf{S}} \sum_{i=1}^{N_{cells}} \sum_{k=1}^{K^{(i)}} \sum_{n \in S_k^{(i)}} P_n^{(i)} = f(\mathbf{P},\mathbf{S}) \tag{4.8}$$

$$r_k^{(i)} \geq R_k^{(i)} \qquad \forall i,k \tag{4.8.1}$$

$$S_j^{(i)} \cap S_k^{(i)} = 0 \qquad \forall k,i,j \neq k \tag{4.8.2}$$

where the constraints are defined in analogy to (4.1), $\mathbf{P}$ is the $N \times N_{cells}$-dimensional vector containing the power allocations in the system, $\mathbf{S}$ is the set of all channel assignments, and $r_k^{(i)} = \sum_{n \in S_k^{(i)}} B \log_2(1+\gamma_{k,n}^{(i)})$ is the transmitted rate of user k in cell i. This optimization problem yields a lower bound on the target problem, to be compared with the value achieved by (4.5).

Unfortunately, problem (4.8) is nonconvex because of the presence of the target rate constraints (4.8.1), which can be expressed as the difference of two concave functions [64], and because of the exclusive allocation constraints (4.8.2). As a consequence, finding the exact solution of the CRA formulation has exponential complexity in N and in the total number of users $N_u = \sum_{i=1}^{N_{cells}} K^{(i)}$ and requires the evaluation of all possible channel assignments. Therefore, rather than following a brute force approach, we resort to finding a (hopefully) close bound to the RRA solution by looking for the maximum of the Lagrange dual function $g(\boldsymbol{\mu})$, which is defined as

$$g(\boldsymbol{\mu}) = \min_{\mathbf{P},\mathbf{S}} \sum_{i=1}^{N_{cells}} \sum_{k=1}^{K^{(i)}} \sum_{n \in S_k^{(i)}} P_n^{(i)} + \sum_{i=1}^{N_{cells}} \sum_{k=1}^{K^{(i)}} \mu_k^{(i)} \left(R_k^{(i)} - r_k^{(i)} \right) \tag{4.9}$$

$$S_j^{(i)} \cap S_k^{(i)} = 0 \qquad \forall k,i,j \neq k \tag{4.9.1}$$

where $\boldsymbol{\mu}$ is the N_u-dimensional vector of Lagrange dual variables (one per user). The solution of the Lagrange dual problem

$$\max g(\boldsymbol{\mu}) \tag{4.10}$$

$$\boldsymbol{\mu} \succeq 0. \tag{4.10.1}$$

yields a lower bound to the solution of the primal problem. Moreover, when the primal problem is convex and satisfies the Slater constraint qualifications, *strong duality* holds, i.e. the gap between the primal solution and the dual solution is zero. In our case the CRA problem is not convex but, under certain conditions [35], strong duality may still hold. In practice, the conditions for strong duality may be difficult to be met, but, since our goal is to assess the quality of the allocation found with LP-based methods, a solution with a small duality gap will suffice.

Unlike (4.8), problem (4.10) is always convex and its solution $\boldsymbol{\mu}^*$ can be found following an iterative strategy based on the ellipsoid [25] method. Although $g(\boldsymbol{\mu})$ is not differentiable due to constraints (4.9.1), we can iteratively update the ellipsoid equation by computing a subgradient [35] of $g(\boldsymbol{\mu})$ as $\mathbf{d} = \mathbf{R} - \mathbf{r}(\mathbf{P}, \mathbf{S})$ where $\mathbf{R}$ and $\mathbf{r}(\mathbf{P}, \mathbf{S})$ are the N_u-dimensional vectors stacking the rate requirements and the rates achieved as function of $\mathbf{P}$ and $\mathbf{S}$ for all the users in the system. Let $\mathbf{P}^{(p)}$ and $\mathbf{S}^{(p)}$ be the allocation vectors computed at iteration p, the recursive algorithm to find $\boldsymbol{\mu}^*$ works as follows:

1. Compute the subgradient vector $\mathbf{d}^{(p+1)} = \mathbf{R} - \mathbf{r}\left(\mathbf{P}^{(p)}, \mathbf{S}^{(p)}\right)$;
2. Given $\mathbf{d}^{(p+1)}$, find $\boldsymbol{\mu}^{(p+1)}$ and $\mathbf{A}^{(p+1)}$ defining the ellipsoid $\mathcal{E}^{(p+1)}$ [25];
3. Given the updated dual variable $\boldsymbol{\mu}^{(p+1)}$, solve (4.9) to compute $\mathbf{P}^{(p+1)}$ and $\mathbf{S}^{(p+1)}$.

Thus, each iteration requires solving (4.9). By assuming that all sub-channels are allocated in all cells, i.e. $\bigcup_{k=1}^{K^{(i)}} S_k^{(i)} = N$ $(i = 1, \ldots, N_{cells})$, and rearranging the order of the sum terms, the problem (4.9) can be reformulated as the sum of N lower-complexity optimization problems

$$g(\boldsymbol{\mu}) = \sum_{n=1}^{N} g_n(\boldsymbol{\mu}) + \sum_{i=1}^{N_{cells}} \sum_{k=1}^{K^{(i)}} \mu_k^{(i)} R_k^{(i)} \tag{4.11}$$

where $g_n(\boldsymbol{\mu})$ is

$$g_n(\boldsymbol{\mu}) = \min_{\mathbf{P},\mathbf{S}} \sum_{i=1}^{N_{cells}} \left(P_n^{(i)} - \sum_{k=1}^{K^{(i)}} \mu_k^{(i)} B \log_2(1 + \gamma_{k,n}^{(i)}) \mathfrak{I}_n\left(S_k^{(i)}\right) \right). \tag{4.12}$$

where the allocation indicator function $\mathfrak{I}_n(S_k)$ for sub-channel n is 1 if $n \in S_k$ and 0 otherwise. Formulation (4.11) allows to split the allocation problem into N distinct unconstrained problems. Unfortunately, since problems (4.12) are nonconvex in $\mathbf{P}$ and $\mathbf{S}$, the complexity, now linear in the number of sub-channels, is still exponential in the number of users. Therefore, following [35] we propose a heuristic that is guaranteed to converge to a local minimum at

least. First of all, we employ a fixed set of M transmission formats $\mathcal{B} = \{0, \beta_1, \beta_2, \ldots, \beta_{M-1}\}$, which corresponds to a set of M target SINRs $\Gamma = \{0, \gamma(1), \gamma(2), \ldots, \gamma(M-1)\}$. Given the dual variable vector $\boldsymbol{\mu}$, we solve (4.12) a cell at the time, i.e. performing the allocation in one cell having fixed the transmission formats and channel assignments in all other cells. In this case, computing the power required to transmit with a given target SINR is a well-known power control problem, which has a closed form solution [65] and, as a consequence, we are able to select the combination of users and formats that minimizes the metric (4.12). This procedure is iterated for each cell, as long as the cells keep changing their allocation. Such an iterative process is guaranteed to converge, because every time a cell changes its allocation the objective function $g_n(\boldsymbol{\mu})$ strictly decreases.

This algorithm is suboptimal in the sense that it may theoretically lead to compute a solution to (4.8), which is only locally optimal. In this case, the vector $\mathbf{d}$ computed using the local optimal rates is not guaranteed to be a subgradient of $g(\boldsymbol{\mu})$. In spite of this, in all our simulations we have never observed any problems in the convergence of the ellipsoid method.

4.4 Numerical Results

We present the numerical results for a multi-cell OFDMA system with $N_{cells} = 7$ hexagonal cells of radius $R = 500$ m. Each BS transmits over a downlink bandwidth $W = 5$ MHz so the sampling time is $T_c = 200$ ns. The path loss is proportional to the distance between the BS and the MS; the path loss exponent is $\alpha = 4$. The overall bandwidth is divided into $N = 16$ sub-channels shared by all the cells in the system. The propagation channel is static and frequency-selective Rayleigh fading with an exponentially decaying power delay profile. The channel delay spread is $\sigma_\tau = 1\ \mu s$ and the coherence bandwidth is assumed to be larger than the bandwidth spanned by each sub-band. Moreover, we consider a population of data users with limited mobility (long channel coherence time). The number of active users at one time is $K^{(i)} = K (i = 1, \ldots, N_{cells})$, with $K = 8$ and, when the scheduler is taken into account, we set $C_{max} = 2N$ and $C_{req} = N$. We assume that all BSs transmit with the same value of spectral efficiency η. The scheduling weights ϕ associated to the users are set to 1, thus the scheduler aims at achieving the same average throughput for all users in a cell. The results shown in the following have been obtained by averaging on 100 channel realizations; each simulation has a duration of 50 frames.

The main performance indicators are P_m, the mean overall transmitted power per cell and

η_m, the measured spectral efficiency averaged over all sub-channels and over all cells; η_m is the average sum rate, normalized to the available bandwidth, and it is computed as:

$$\eta_m = \eta E\left\{C_{req}\left(\eta\right)\right\}/N, \tag{4.13}$$

where $C_{req}(\eta)$ is the cell load for a given value of η.

The number of iterations of the allocation phase is the sum of N_s and N_{pow}. In practice $N_{pow} = 3$ iterations are already sufficient for power control, so that most of the iterations are due to the effort of finding a stable allocation. Thus N_s is a critical parameter: if it is too small, the allocation may not have enough time to converge and an excessive number of sub-channels may have to be switched off. However, increasing N_s leads to more overhead. In practice, relatively small values of N_s are sufficient to achieve a high spectral efficiency in the transmission phase.

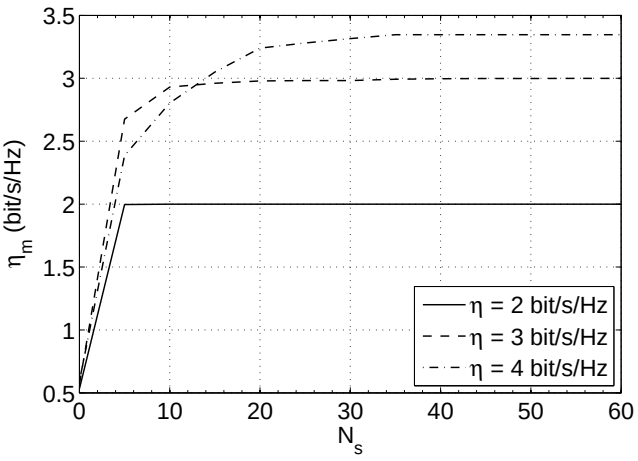

Figure 4.2: Measured spectral efficiency η_m vs. duration of allocation phase N_s

Figures 4.2 and 4.3 illustrate how our DLA performs as a function of N_s for various values of η. For intermediate values of the load ($\eta = 2$ bit/s/Hz) a few iterations are sufficient to find a stable allocation with no need to reduce the cell load. As the load grows ($\eta = 3$ bit/s/Hz), the DLA needs a larger number of iterations to find a feasible allocation but the system still performs very little load reduction. Further increases of the load ($\eta = 4$ bit/s/Hz) require a much larger number of iterations and impose substantial load reductions. It is worth comparing the performance for $\eta = 3$ and $\eta = 4$ bit/s/Hz after 10 iterations: the system measures approximately the same mean spectral efficiency but for $\eta = 4$ consumes almost twice

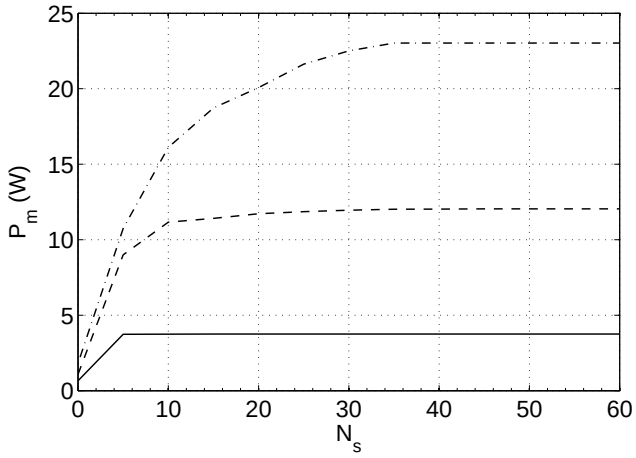

Figure 4.3: Mean power per cell P_m vs. duration of allocation phase N_s

as much power. In the following we set $N_s = 5$ for $\eta \leq 2$, $N_s = 10$ for $2 < \eta \leq 3$ and $N_s = 20$ for $3 < \eta \leq 4$ bit/s/Hz.

In the following (Figures 4.6 and 4.5), we compare the performance of the LPRA with other RA schemes encountered in the literature. Since some of the schemes analyzed do not integrate easily in our layered architecture, the results are obtained simulating only one frame. In this case, we impose the same rate requirements to all users in the system and we do not take into account the PS. As a consequence, in case of load reduction, any fairness issue is neglected.

The MARA [54] is an iterative heuristic based on a LP approach. Starting from the formulation (4.5), each iteration of the MARA is composed of two phases. In the first phase, channels are allocated neglecting the MAI so that the power costs are computed based on the channel gains with the serving BS. In the second phase the allocator deals with interference by solving a centralized power control problem on each sub-channel. At the end of each iteration the algorithm dissuades the users from using the sub-channels where they cannot achieve their target SINR by artificially increasing the cost of those specific sub-channels. To enforce convergence, the rate constraints for those users that are unable to find a stable allocation are progressively reduced.

The scheme proposed by Pischella and Belfiore (PBRA) in [55] is an iterative algorithm made up of several steps: 1) determine for each user and each sub-channel a maximum SINR value so that the convergence of distributed power control is guaranteed; 2) allocate sub-channels to the users according to a heuristic that aims at power minimization under target data rate

requirements; 3) solve a convex optimization problem that sets a SINR target for each user respecting the constraints defined in step 1) and meeting the user rate constraints; 4) perform distributed power control to meet the SINR targets of step 3). Users for which it is not possible to meet the rate constraints in step 3) are simply switched off.

We also compare the system performance with the two FFR schemes presented in [56], denoted as FFR-A and FFR-B, which enforce different allocation strategies on the base of the position of the users in the cell. In the FFR-A scheme terminals close to the BS use the same bandwidth in all the cells, while the channels assigned to edge users can not be reused in adjacent cells. The FFR-B strategy allows users near the BS and edge users located in adjacent cells to transmit on the same channels. Given a population of users, these FFR schemes are employed to build an interference graph where two MSs are connected by an edge if they interfere with each other. The resources are then assigned following a greedy strategy based on the channel gains with the serving BS, with the objective of minimizing the overall power subject to interference and rate constraints. Once all channels are assigned, each BS performs N_{pow} power control iterations and those sub-channels which have not yet achieved their target SINR are switched off.

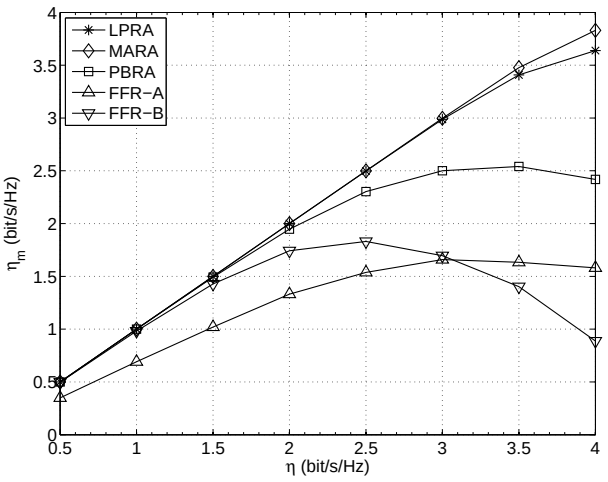

Figure 4.4: Measured spectral efficiency η_m vs. target spectral efficiency η (resource allocation and load control, no scheduler, same rate for all users)

Figure 4.6 plots η_m versus η for the different allocators. Since FFR-A's interference graph has many connections, even at low traffic loads it does not manage to satisfy all user requirements; for this reason FFR-B and all the other allocators outperform it. On the other hand, for high

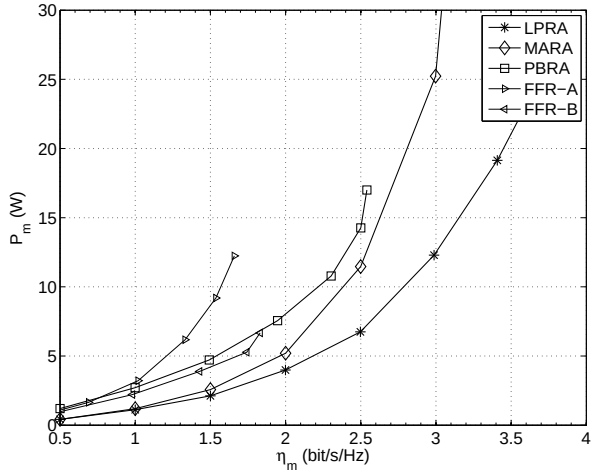

Figure 4.5: Mean power per cell P_m vs. measured spectral efficiency η_m (resource allocation and load control, no scheduler, same rate for all users)

traffic loads FFR-A is more robust to interference and performs better than FFR-B. For $\eta > 2$ bit/s/Hz, the PBRA algorithm achieves a stable allocation by significantly reducing the load, while both LPRA and MARA schemes require a much more limited reduction in throughput. For $\eta > 3$ bit/s/Hz, the centralized heuristic MARA is only slightly better than our solution.

Figure 4.5 plots P_m vs η_m. For a given value of η_m LPRA consumes the least power of all studied schemes. MARA has very good performance with low to medium loads but for high loads it is unable to cope with the increasingly strong MAI. FFR-B and PBRA have similar results while FFR-A is penalized by the diversity, consequence of its interference management policy.

A further insight into the performance of the distributed iterative LPRA scheme is obtained by the comparison with the performance of the CRA algorithm presented in Section 4.3.

Figure 4.7 shows the mean transmitted power P_m as a function of K, the number of users per cell, for LPRA with $\eta = 2$ and CRA with $M = 2$ and $M = 4$. M denotes the number of transmission formats; specifically, with $M = 2$ the two formats correspond to spectral efficiency $\eta = 0$ (no transmission) and $\eta = 2$, with $M = 4$ the four formats correspond to $\eta = 0, 1, 2, 4$. For a small number of users CRA with $M = 4$ significantly outperforms the other schemes and the reason thereof is that link adaptation is very important when there is little multi-user diversity. As the number of users increases, the performance gap between CRA with $M = 4$ and CRA with $M = 2$ transmission format and LPRA becomes almost negligible. Two important

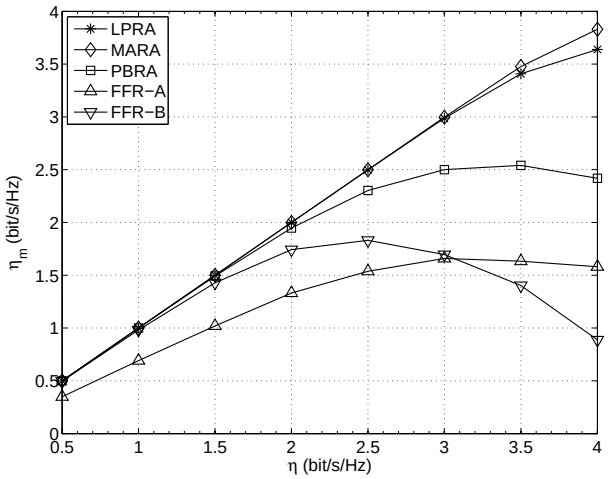

Figure 4.6: Measured spectral efficiency η_m vs. target spectral efficiency η (resource allocation and load control, no scheduler, same rate for all users)

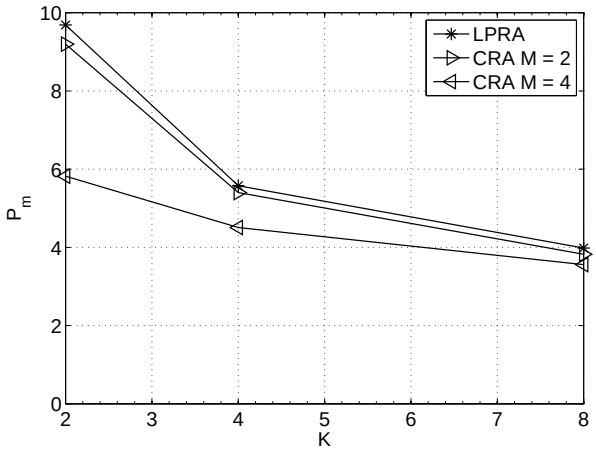

Figure 4.7: Mean power per cell P_m vs. number of users K (resource allocation and load control, no scheduler, same rate for all users)

conclusions can be drawn by the analysis of the results in Figure 4.7: i) there is a very small penalty for adopting just one transmission format, provided that there are enough users in the system; ii) in the scenarios under study the information acquired by feedbacking the interference in an iterative distributed algorithm makes it possible to achieve results that are close to those obtained with centralized algorithms possessing full channel state information for all users in

the multi-cell system.

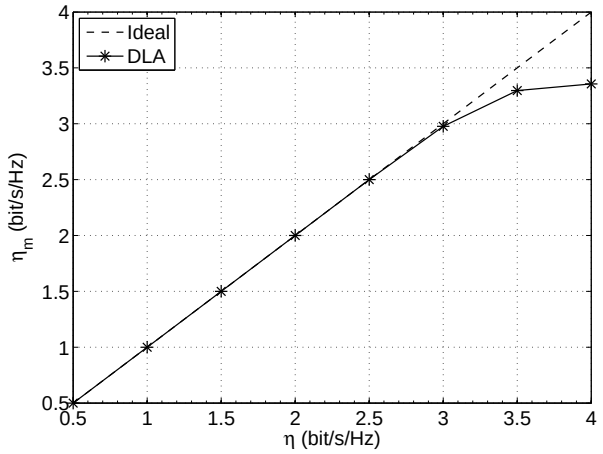

Figure 4.8: Measured spectral efficiency η_m vs. target spectral efficiency (Distributed Layered Architecture)

Figure 4.8 shows the performance of the full-blown DLA as function of η. The difference with respect to the results presented in Figure 4.6 is that now the scheduler dictates, frame by frame, the rate constraints of the users, in order to achieve fairness in the long term. The dashed line plots the curve of the ideal case when no load reduction takes place, i.e. $\eta_m = \eta$: the results are very close even when fairness is actively pursued. In general, we can show that the DLA is able to cope with interference even for high traffic loads.

A commonly used measure of fairness for the throughputs x_i of a set of n flows is Jain's fairness index, defined as $\left(\sum_{i=1}^{n} x_i\right)^2 / \left(n \sum_{i=1}^{n} x_i^2\right)$. Figure 4.9 plots Jain's fairness index averaged among cells for the three values of spectral efficiency $\eta = 2, 3, 4$ bit/s/Hz. It is worth pointing out that, even in high load conditions (i.e. at $\eta = 4$ bit/s/Hz), the packet scheduler is able to guarantee a fair throughput allocation among users in the cells in the long term. From Figure 4.9, it appears that a fairness index in excess of 0.99 is achieved in at most 20 frames. For a frame duration in the order of tens of ms this amounts to a few hundred ms, which implies that fair throughput is perceived from the applications' point of view.

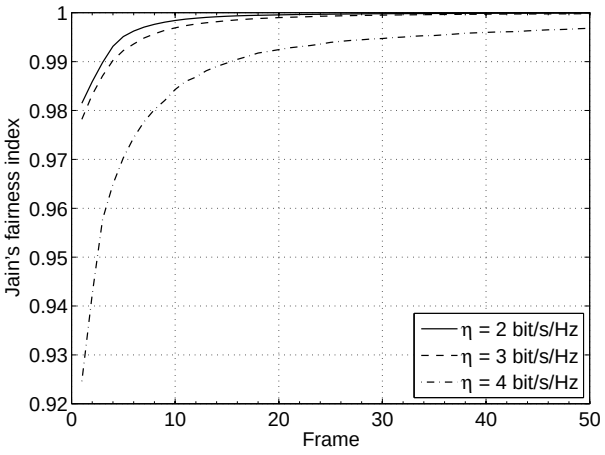

Figure 4.9: Jain's fairness index per cell vs. frame

Chapter 5

Carrier frequency offset recovery in FBMC systems

We already discussed about the main characteristics of FBMC modulation and its appealing features that led it to be one of the most important technologies adopted in several standards for next generation multicarrier systems. The increasing interest in FBMC modulation is also demonstrated by a number of relevant contributions on the topics of channel equalization and signal synchronization. In [66], per-subcarrier equalization is performed by employing a first-order allpass and linear-phase FIR filter, while in [67] fractionally spaced linear and decision feedback equalizers for FBMC are designed and analyzed. The timing synchronization problem is addressed in [68]-[71]. In detail, the BER sensitivity to timing errors is discussed in [68], whereas in [69] and [70] simple data-aided and decision directed timing error detectors. In [71], instead, an iterative blind closed-loop scheme is derived. The issue of CFO recovery for multicarrier systems has received considerable attention in recent literature, primarily for application in the OFDM context [72]-[77], but unfortunately the techniques proposed for OFDM are neither optimized nor directly applicable to FBMC in view of the rather different signal formats.

However, a few algorithms for CFO extraction specifically tailored for FBMC have also been proposed, and they can be roughly qualified as PA or blind schemes, depending on whether they rely or not on the presence of known pilot symbols multiplexed with the data stream. A PA estimator attempts to recover the CFO (and possibly other parameters of interest such as channel state information, timing error etc.) from the received samples associated to the

pilot symbols after modulation removal [78], [79], while blind schemes extract CFO information from manipulation of proper symbol-independent metrics, notably second-order statistics such as the energy collected at the matched filter output [13] or specific correlation properties of the multicarrier waveform [80], [81].

The above references provide a representative sample of the scarce literature available on CFO recovery for FBMC. All of them, however, with the exception of [78], propose algorithms devised under the assumption of time-invariant or slowly changing fading and as such they exhibit a poor behavior whenever the channel is affected by a significant Doppler spread, in addition to frequency selectivity, a condition frequently met in wireless mobile communications.

In [78] this aspect is specifically addressed and the double (time-frequency) variability of the channel is dealt with by incorporating the statistical knowledge of the channel, in terms of autocovariance matrix of the channel fading, in a ML-derived estimation algorithm. The cited matrix is evaluated at the pilot positions, uniformly spaced apart both in frequency (over the available subcarriers) and in time (over the symbol positions). A major issue with this approach is related to the correct choice of the pilot symbol spacing, that cannot be too tight not to impact spectrum efficiency, nor too loose to permit correct sampling of the time-frequency fading process, to be pursued not only in view of CFO recovery but also for accurate channel estimation and equalization. When the pilot pattern is properly selected, a PA-based MLE (PA-MLE) can normally boast a smaller estimation error in comparison with blind counterparts [78]. On the other hand, the range of values of CFO ensuring unambiguous operation of the PA-MLE (i.e., its *acquisition range*) turns out to be approximately as wide as the inverse of the time-domain pilot separation, and hence may become very narrow when the latter are widely spaced apart, a limitation that is not felt when a blind approach is employed. The above may represent a serious limitation in those situations where the CFO is a significant fraction of the subcarrier symbol rate, as occurs in high-data-rate applications or in the presence of a large relative instability of the inherent transmission-reception oscillators.

In this Chapter we make a step ahead with respect to the approach in [78] by proposing an improved version with extended acquisition while still ensuring adequate accuracy at affordable complexity. We derive a CFO recovery algorithm that takes advantage both of pilot symbols and also indirectly of data symbols through knowledge and exploitation of their modulation format. More specifically, in addition to the information provided by the pilot symbols, the new algorithm attempts to improve the CFO estimate by utilizing the phase rotation introduced by the CFO between any other available pair of symbols, either of mixed type (i.e., a pilot and a

data symbol) or homologous type (two data symbols), through previous coarse estimation and cancellation of the relative (differential) phase shift induced by the modulation. The latter shift belongs to the finite-size set of all possible differential phases between constellation symbols, and therefore it can be reliably estimated from the two observed samples provided that the random phase drift due to the combined effect of noise, CFO and channel fading can be considered small with respect to the minimum angular separation between constellation points. This constraint however is not as restrictive as it might appear at first glance since, as discussed in the following, the modulation-induced differential phase between any pair of received symbols can be decomposed into a sequence of concatenated phase shifts between consecutive symbols, and therefore the cited requirement turns out to be relatively mild as it applies ultimately to a single symbol interval.

The rationale behind the proposed approach can be briefly outlined as follows. Initially the algorithm is set out in the framework of ML joint estimation of CFO and data symbols. For simplicity only the 4-QAM modulation format is considered. Subsequent elaboration shows that the LLF depends on the differential phases between pairs of tentative symbols in addition to the CFO. Next, the dependence of the LLF on the data symbols is dropped by replacing the above differential phases by their estimates built, as mentioned, from the concatenation of estimates relevant to pairs of neighboring samples. Finally, a search for the maximum of the resulting LLF leads to the desired estimate of the CFO.

5.1 Signal Model

As shown in Chapter 2, the FBMC signal transmitted over a burst of L consecutive blocks can be written as

$$s(t) = \sum_{\ell=1}^{L} \sum_{q=1}^{N} a_{\ell}^{(q)} g(t - \ell T) e^{j2\pi q M t/NT}, \tag{5.1}$$

where $a_{\ell}^{(q)}$ is the (information or pilot) symbol transmitted over the $l-$th FBMC block within the $q-$th subcarrier, $1 \leq q \leq N$, and constraining the quantity $M \triangleq (1+\xi)N$ to be an integer for ease of implementation. Under some mild assumptions (see 2.2.2), the received sample taken on the $n-$th subcarrier at the instant $t = kT$ can be approximated as

$$z_k^{(n)} \simeq \varphi_k^{(n)} a_k^{(n)} e^{j2\pi\nu k} + w_k^{(n)}, \tag{5.2}$$

where $\varphi_k^{(n)}$ is a multiplicative factor accounting for the channel time-frequency selectivity, ν is the CFO normalized to the signaling rate $1/T$, ϑ is the phase offset, and $w_k^{(n)}$ is a complex-valued

zero-mean independent Gaussian random variable with variance $\sigma^2 = 2N_0$.

Let us define

$$\mathbf{z}^{(n)} \triangleq [z_1^{(n)}, \ldots, z_L^{(n)}]^T, \tag{5.3}$$

$$\boldsymbol{\varphi}^{(n)} \triangleq [\varphi_1^{(n)}, \ldots, \varphi_L^{(n)}]^T, \tag{5.4}$$

$$\mathbf{w}^{(n)} \triangleq [w_1^{(n)}, \ldots, w_L^{(n)}]^T \tag{5.5}$$

as the vectors collecting the received samples, the time-varying channel gains and the noise samples, respectively, on the $n-$th subcarrier. Thus, from (5.2) it follows that

$$\mathbf{z}^{(n)} = \boldsymbol{\Psi}(\nu)\mathbf{A}^{(n)}\boldsymbol{\varphi}^{(n)} + \mathbf{w}^{(n)}, \tag{5.6}$$

where $\boldsymbol{\Psi}(\nu)$ and $\mathbf{A}^{(n)}$ are $L \times L$ diagonal matrices defined as

$$\boldsymbol{\Psi}(\nu) \triangleq \text{diag}\left\{e^{j2\pi\nu}, \ldots, e^{j2\pi\nu L}\right\}, \tag{5.7}$$

and

$$\mathbf{A}^{(n)} \triangleq \text{diag}\left\{a_1^{(n)}, \ldots, a_L^{(n)}\right\}. \tag{5.8}$$

Hence, stacking $\mathbf{z}^{(n)}$, $\boldsymbol{\varphi}^{(n)}$ and $\mathbf{w}^{(n)}$ to form the vectors

$$\mathbf{z} \triangleq [\mathbf{z}^{(1)^T}, \ldots, \mathbf{z}^{(N)^T}]^T, \tag{5.9}$$

$$\boldsymbol{\varphi} \triangleq [\boldsymbol{\varphi}^{(1)^T}, \ldots, \boldsymbol{\varphi}^{(N)^T}]^T, \tag{5.10}$$

$$\mathbf{w} \triangleq [\mathbf{w}^{(1)^T}, \ldots, \mathbf{w}^{(N)^T}]^T, \tag{5.11}$$

the observed sequence can be expressed as

$$\mathbf{z} = \boldsymbol{\Gamma}(\nu)\mathbf{A}\boldsymbol{\varphi} + \mathbf{w}, \tag{5.12}$$

where $\boldsymbol{\Gamma}(\nu)$ and $\mathbf{A}$ are $LN \times LN$ diagonal matrices defined as

$$\boldsymbol{\Gamma}(\nu) \triangleq \text{Diag}\left\{\boldsymbol{\Psi}(\nu), \ldots, \boldsymbol{\Psi}(\nu)\right\}, \tag{5.13}$$

and

$$\mathbf{A} \triangleq \text{Diag}\left\{\mathbf{A}^{(1)}, \ldots, \mathbf{A}^{(N)}\right\}. \tag{5.14}$$

5.2 ML Carrier Frequency Estimation

5.2.1 Formulation of the ML Estimation Problem

In a typical wireless propagation environment, the set of channel coefficients $\boldsymbol{\varphi}$ can be modeled as a complex zero-mean Gaussian vector with covariance matrix $\mathbf{C}_{\boldsymbol{\varphi}} = E\{\boldsymbol{\varphi}\boldsymbol{\varphi}^H\}$. Since the latter changes slowly with time, it can be estimated at the receiver either through channel sounding or exploiting the a priori knowledge about the multipath delay profile and the Doppler spread. Under this assumption, the received samples $\mathbf{z}$ in (5.12), for given CFO ν and $\mathbf{A}$, are jointly complex zero-mean Gaussian as well, with covariance matrix $\mathbf{C}_{\mathbf{z}}(\nu) = E\{\mathbf{z}\mathbf{z}^H\} = \boldsymbol{\Gamma}(\nu)\mathbf{A}\mathbf{C}_{\boldsymbol{\varphi}}\mathbf{A}^H\boldsymbol{\Gamma}(\nu)^H + \sigma^2\mathbf{I}_{LN}$. Therefore, the JMLE of the normalized CFO and the transmitted symbols is obtained by minimizing the LLF [82]

$$\Lambda_{\mathrm{JMLE}}(\nu, \mathbf{A}) \triangleq \mathbf{z}^H\boldsymbol{\Gamma}(\nu)\mathbf{A}\mathbf{F}\mathbf{A}^H\boldsymbol{\Gamma}(\nu)^H\mathbf{z}, \tag{5.15}$$

where $\mathbf{F} \triangleq \left(\mathbf{C}_{\boldsymbol{\varphi}} + \sigma^2\mathbf{I}_{LN}\right)^{-1}$. Finding the JMLE solution involves, however, a linear grid search over the interval $\mathcal{I}$ including the possible values of ν together with all possible realizations of $\mathbf{A}$. Due to its combinatorial nature, the computational complexity of this problem becomes intractable even for small L and N, and consequently, alternative methods for CFO recovery need to be devised.

A good complexity versus accuracy tradeoff is offered by the PA-MLE scheme pursued in [78]. Herein, the main assumption is that each of S subcarriers conveys P pilot symbols, with $S \leq N$, $P \leq L$, so that the total number of pilots embedded in the burst amounts to $Q \triangleq PS$. Bearing in mind that the observation model (5.12) still holds when applied to the subset of the received samples corresponding to pilot positions (with the difference that $\mathbf{z}$, $\boldsymbol{\varphi}$ and $\mathbf{w}$ are Q-dimensional vectors and $\boldsymbol{\Gamma}(\nu)$, $\mathbf{A}$ and $\mathbf{C}_{\boldsymbol{\varphi}}$ are $Q \times Q$ diagonal matrices), we can build an LLF metric similar to (5.15), that depends no longer on the (unknown) data symbols but only on the frequency offset to be estimated; see [78] for further details. The following aspects concerning the minimization of the LLF for the PA-MLE are now to be noted.

1. Since the matrix inversion required in the LLF can be performed off-line, only a linear search is required over the interval $\mathcal{I}$ of trial values for ν (coarse search), followed by interpolation (fine search).

2. It is convenient to remove the dependence of $\mathbf{F}$ on σ by setting it to a predefined value σ_0, e.g. the one corresponding to the SNR at the nominal receiver sensitivity.

3. It can be argued that the acquisition range of the PA-MLE widens by reducing the time-domain pilot spacing, i.e., by increasing the number of symbols Q. At the same time, the pilot overhead $\chi \triangleq \frac{Q}{NL}$ (and therefore Q for a given NL) has to be chosen as low as possible, say below 10 %, to not degrade the power and spectral system efficiency.

5.2.2 Combined Pilot-Aided and Decision-Directed Frequency Estimation

The above contrasting requirements about the choice of Q motivate the search for a more efficient alternative to the PA-MLE scheme. Rather than increasing Q, i.e., the pilot overhead χ, we pursue a different estimation strategy, based on combining the PA and DD concepts together, that in the sequel will be referred to as CMLE.

The LLF in (5.15) for the JMLE of ν and $\mathbf{A}$ can be rearranged into

$$\begin{aligned}\Lambda_{\mathrm{JMLE}}(\nu,\mathbf{A}) = &\sum_{l=1}^{L}\sum_{s,c=1}^{N} a_l^{(s)} a_l^{(c)^*} z_l^{(s)^*} z_l^{(c)} G_{l,l}^{(s,c)} \\ &+ 2\sum_{k=1}^{L-1}\sum_{l=1}^{L-k} \mathrm{Re}\left\{ e^{-j2\pi\nu k} \sum_{s,c=1}^{N} a_l^{(s)} a_{l+k}^{(c)^*} z_l^{(s)^*} z_{l+k}^{(c)} G_{l,l+k}^{(s,c)} \right\},\end{aligned} \tag{5.16}$$

where $G_{n,m}^{(s,c)} \triangleq [\mathbf{G}]_{(s-1)L+n,(c-1)L+m}$, with $\mathbf{G}$ being defined as the matrix $\mathbf{F}$ in (5.15) evaluated at $\sigma = \sigma_0$, i.e., $\mathbf{G} \triangleq \left(\mathbf{C}_{\boldsymbol{\varphi}} + \sigma_0^2 \mathbf{I}_{LN}\right)^{-1}$. The LLF (5.16) depends on all the products between any two symbols (data or pilots) within the burst. This fact suggests that we can get rid of data symbols by resorting to differential decisions taken on the products $z_l^{(s)} z_{l+k}^{(c)^*}$. To be specific, from (5.2) we get

$$z_l^{(s)} z_{l+k}^{(c)^*} = \varphi_l^{(s)} \varphi_{l+k}^{(c)^*} a_l^{(s)} a_{l+k}^{(c)^*} e^{-j2\pi\nu k} + \zeta_{l,l+k}^{(s,c)}, \tag{5.17}$$

where $\zeta_{l,l+k}^{(s,c)}$ is a disturbance term accounting for both the channel noise and fading. About (5.17), we invoke the following assumptions:

a1) the channel fading is sufficiently correlated in both time- and frequency domains so that $\varphi_l^{(s)} \varphi_{l+k}^{(c)^*} \simeq \left|\varphi_l^{(s)}\right|^2$, regardless of the time lag k and the subcarrier indexes s and c;

a2) the normalized CFO ν is sufficiently small so that $e^{-j2\pi\nu k} \simeq 1$, $\forall k \in [1, L]$;

a3) the data and pilot symbols belong to a 4-QAM constellation with unitary radius, and all pilots are equal.

Next, after defining the *differential symbol* $b_{l,l+k}^{(s,c)} \triangleq a_l^{(s)} a_{l+k}^{(c)^*}$, we apply a1)-a3) in (5.17) yielding

$$z_l^{(s)} z_{l+k}^{(c)^*} \simeq \left|\varphi_l^{(s)}\right|^2 b_{l,l+k}^{(s,c)} + \zeta_{l,l+k}^{(s,c)}. \tag{5.18}$$

Hence, the differential symbol $b_{l,l+k}^{(s,c)}$, that in view of a3) belongs to a 4-QAM constellation, can be estimated, independently of the knowledge of the CFO and the fading gains, by feeding the product $z_l^{(s)} z_{l+k}^{(c)^*}$ into a conventional threshold detector. In order to give strength to assumptions a1) and a2) for all values of k and s, c, with $s \neq c$, we rearrange the differential symbol as $b_{l,q}^{(s,c)} = a_l^{(s)} p_{\bar{l}}^{(\bar{s})^*} p_{\bar{q}}^{(\bar{c})} a_q^{(c)^*} = b_{l,\bar{l}}^{(s,\bar{s})} b_{\bar{q},q}^{(\bar{c},c)}$, where $p_{\bar{l}}^{(\bar{s})}$ and $p_{\bar{q}}^{(\bar{c})}$ are the pilots closest to the data symbols $a_l^{(s)}$ and $a_q^{(c)}$, respectively. The result is that the differential decision can be computed as $\hat{b}_{l,q}^{(s,c)} = \hat{b}_{l,\bar{l}}^{(s,\bar{s})} \hat{b}_{\bar{q},q}^{(\bar{c},c)}$, i.e., as the product of two differential decisions on two closely spaced symbols, and therefore, error propagation can be strongly mitigated. Thus, replacing the product $a_l^{(s)} a_{l+k}^{(c)^*}$ in (5.16) by the differential decision $\hat{b}_{l,l+k}^{(s,c)}$ and dropping immaterial factors, we obtain the modified metric

$$\Lambda_{\mathrm{CMLE}}(\nu) = \sum_{k=1}^{L-1} \sum_{l=1}^{L-k} \mathrm{Re}\left\{ e^{-j2\pi\nu k} \sum_{s,c=1}^{N} \hat{b}_{l,l+k}^{(s,c)} z_l^{(s)^*} z_{l+k}^{(c)} G_{l,l+k}^{(s,c)} \right\}, \tag{5.19}$$

whose minimization provides the CMLE solution.

We now focus on some issues concerning the computational complexity of (5.19).

1. Compared to (5.16), the metric (5.19) does no longer depend on data symbols, since the latter have been replaced by differential decisions as discussed earlier. This means that the CMLE requires a linear grid search over the interval $\mathcal{J}$ as well, through the same two-step procedure already discussed for the PA-MLE.

2. Due to the specific structure of the fading vector $\boldsymbol{\varphi}$, the covariance $\mathbf{C}_{\boldsymbol{\varphi}}$ is a block-Toeplitz matrix composed of N^2 sub-matrices $\mathbf{C}_{\boldsymbol{\varphi}}^{(n,m)}$ of size $L \times L$, $1 \leq n, m \leq N$, representing the covariance between the subcarriers of indexes n and m. In order to drastically simplify the CMLE algorithm, the correlation between the fading coefficients belonging to different subcarriers has been dropped, thus replacing $\mathbf{C}_{\boldsymbol{\varphi}}$ by $\bar{\mathbf{C}}_{\boldsymbol{\varphi}} \triangleq \mathrm{Diag}\left\{\mathbf{C}_{\boldsymbol{\varphi}}^{(1,1)}, \ldots, \mathbf{C}_{\boldsymbol{\varphi}}^{(N,N)}\right\}$. As a consequence, the simplified CMLE employing the covariance $\bar{\mathbf{C}}_{\boldsymbol{\varphi}}$ is mismatched to the actual fading conditions. Nevertheless, the resultant performance loss is moderate, as as is proved by simulation results in Sect. 5.4. Thanks to the fading stationarity property, however, we have $\bar{\mathbf{C}}_{\boldsymbol{\varphi}} = \mathrm{Diag}\{\boldsymbol{\Omega}, \ldots, \boldsymbol{\Omega}\}$, with $\boldsymbol{\Omega} \triangleq \mathbf{C}_{\boldsymbol{\varphi}}^{(1,1)} = \mathbf{C}_{\boldsymbol{\varphi}}^{(2,2)} = \cdots = \mathbf{C}_{\boldsymbol{\varphi}}^{(N,N)}$, and therefore, the inverse $\bar{\mathbf{G}} \triangleq \left(\bar{\mathbf{C}}_{\boldsymbol{\varphi}} + \sigma_0^2 \mathbf{I}_{LN}\right)^{-1}$ can be easily computed as a block diagonal matrix whose blocks are all equal to $\left(\boldsymbol{\Omega} + \sigma_0^2 \mathbf{I}_L\right)^{-1}$. The CFO recovery algorithm employing $\bar{\mathbf{G}}$ instead of $\mathbf{G}$ in the metric (5.19) will be designated as LC-CMLE.

3. To quantify the complexity level of the CFO recovery schemes illustrated so far, let us take the subsequent assumptions: *i*) the computational load is provided by the number of real-valued operations (additions and multiplications) required to compute the metric to be optimized for each of the N_ν (equispaced) trial values within the interval $\mathcal{I}$; *ii*) given $\mathbf{C}_{\boldsymbol{\varphi}}$ and σ_0, the matrix $\mathbf{G}$ is computed only once as product between an upper and lower triangular matrix (Cholesky decomposition); *iii*) the products involved in the differential decisions are not taken into account, since they correspond to phase rotations by multiples of $\pi/2$ and can be performed by swapping the real and/or imaginary parts of the operands; *iv*) the values of $e^{-j2\pi\nu k}$ are stored in a look-up table; *v*) the parameters N, L and N_ν are of the same order of magnitude. Without delving into details, it can be shown that the overall complexity requirements of the PA-MLE, CMLE and LC-CMLE schemes are given by $\mathcal{C}_{\text{PA-MLE}} = \mathcal{O}(N_\nu Q^2)$, $\mathcal{C}_{\text{CMLE}} = \mathcal{O}(N^2L^2)$ and $\mathcal{C}_{\text{LC-CMLE}} = \mathcal{O}(NL^2)$, respectively. Therefore, it comes out that the LC-CMLE, when compared to the CMLE, has a complexity reduced by a factor of N, whereas it is approximately as complex as the PA-MLE, even though the latter depends on the number N_ν of the CFO trial values.

5.3 Cramer-Rao Lower Bound for PA-MLE

The CRLB for the CFO estimate is derived under the assumption that all the symbols transmitted within the burst (included in $\mathbf{A}$) are known. Bearing in mind that the PDF $p(\mathbf{z};\nu)$ of the received samples $\mathbf{z}$ in (5.12) for a given CFO ν is complex Gaussian with zero-mean and covariance matrix $\mathbf{C}_{\mathbf{z}}(\nu)$, the Fisher information is given by [82]

$$I(\nu) = -\mathrm{E}\left\{\frac{\partial^2 \ln p(\mathbf{z};\nu)}{\partial \nu^2}\right\} = \mathrm{tr}\left\{\left[\mathbf{C}_{\mathbf{z}}^{-1}(\nu)\frac{\partial \mathbf{C}_{\mathbf{z}}(\nu)}{\partial \nu}\right]^2\right\}, \tag{5.20}$$

and correspondingly, $\mathrm{CRLB}(\nu) = I^{-1}(\nu)$.

Putting the covariance matrix of $\mathbf{z}$ in the form $\mathbf{C}_{\mathbf{z}}(\nu) = \boldsymbol{\Gamma}(\nu)(\mathbf{A}\mathbf{C}_{\boldsymbol{\varphi}}\mathbf{A}^H + \sigma^2\mathbf{I}_{LN})\boldsymbol{\Gamma}(\nu)^H$, its derivative with respect to ν is found to be

$$\frac{\partial \mathbf{C}_{\mathbf{z}}(\nu)}{\partial \nu} = j2\pi\boldsymbol{\Gamma}(\nu)\left[\mathbf{H}(\mathbf{A}\mathbf{C}_{\boldsymbol{\varphi}}\mathbf{A}^H + \sigma^2\mathbf{I}_{LN}) - (\mathbf{A}\mathbf{C}_{\boldsymbol{\varphi}}\mathbf{A}^H + \sigma^2\mathbf{I}_{LN})\mathbf{H}\right]\boldsymbol{\Gamma}(\nu)^H, \tag{5.21}$$

where the $LN \times LN$ diagonal matrix $\mathbf{H}$ is defined as

$$\mathbf{H} \triangleq \mathrm{diag}\left\{1,\ldots,L,\ldots,1,\ldots,L\right\}. \tag{5.22}$$

Hence, replacing the inverse matrix $\mathbf{C}_{\mathbf{z}}^{-1}(\nu) = \boldsymbol{\Gamma}(\nu)(\mathbf{A}\mathbf{C}_{\boldsymbol{\varphi}}\mathbf{A}^H + \sigma^2\mathbf{I}_{LN})^{-1}\boldsymbol{\Gamma}(\nu)^H$ and (5.21) in

(5.20) produces after some passages the desired result

$$I(\nu) = 8\pi^2 \mathrm{tr}\left\{(\mathbf{C}_{\boldsymbol{\varphi}} + \sigma^2 \mathbf{I}_{LN})^{-1} \mathbf{H} (\mathbf{C}_{\boldsymbol{\varphi}} + \sigma^2 \mathbf{I}_{LN}) \mathbf{H} - \mathbf{H}^2\right\}, \tag{5.23}$$

with σ^2 being the inverse of the signal-to-noise (SNR) ratio E_s/N_0. Finally, evaluating (5.23) for $\sigma \to \infty$ (low-SNR) yields

$$I_\infty(\nu) = 8\pi^2 \left(\frac{E_s}{N_0}\right)^2 \mathrm{tr}\left\{\mathbf{C}_{\boldsymbol{\varphi}}^2 \mathbf{H}^2 - (\mathbf{C}_{\boldsymbol{\varphi}} \mathbf{H})^2\right\}, \tag{5.24}$$

whereas for $\sigma \to 0$ (high-SNR) we obtain

$$I_0(\nu) = 8\pi^2 \mathrm{tr}\left\{\mathbf{C}_{\boldsymbol{\varphi}}^{-1} \mathbf{H} \mathbf{C}_{\boldsymbol{\varphi}} \mathbf{H} - \mathbf{H}^2\right\}, \tag{5.25}$$

which proves the existence of a floor in the CRLB due to the time-variance of the fading channel.

5.4 Performance Results

In this section, the performance of the PA-MLE, CMLE and LC-CMLE estimators is evaluated through computer simulation. These results are compared against two benchmarks: one being the PA-MLE scheme with $Q = NL$, i.e., where all the burst symbols are assumed known, which will be referred to as DA-MLE, the other the corresponding CRLB, as derived in Section . First, we concentrate on the mean square estimation error (MSEE) $\mathrm{E}\{(\hat{\nu} - \nu)^2\}$ as a function of the true CFO ν, that is useful to identify the *acquisition range* where the estimator performance is reliable. Next, the *noise sensitivity* is analyzed for a specified CFO in terms of the MSEE as a function of the mean-energy-per-symbol-to-noise-spectral-density ratio E_s/N_0.

5.4.1 Simulation Setup

Each burst is composed of $N = 16$ subcarriers and spans $L = 50$ FBMC symbol intervals. As case study (other pilot patterns have been tested), we choose the pilot overhead $\chi = 5\,\%$, so that the total number of pilots is $Q = 40$, and $S = N/2 = 8$ subcarriers each bearing $P = 5$ pilots, with $\mathcal{N} = \{0, 2, 4, 6, 9, 11, 13, 15\}$ and $\mathcal{K} = \{0, 9, 19, 29, 39\}$ being the pilot indexes along the time and frequency axes, respectively. The SRRC prototype filter $g(t)$ has roll-off factor $\xi = 0.25$, while both data and pilot symbols are unitary-energy with 4-QAM format. The channel is time- and frequency-selective with $U = 6$ paths. The excess delays $\Delta\tau_u \triangleq \tau_u - \tau_0$ are uniformly spaced, and the gains are modelled as zero-mean independent complex-valued Gaussian processes having Jakes power spectrum with Doppler bandwidth f_D and mean

square value decaying exponentially with $\Delta\tau_u$, so that the normalized channel delay spread is $\sigma_\tau/T = 10^{-2}$. Two different propagation environments are considered: SF channels with $f_D T = 7.5 \cdot 10^{-3}$ and FF channels with $f_D T = 3 \cdot 10^{-2}$. As for the noise variance required by the estimators, a single predefined value of σ_0^2 is selected so that the corresponding E_s/N_0 is equal to 15 dB.

5.4.2 MSEE Performance

Acquisition Range. The acquisition range evaluated at E_s/N_0 of 30 dB is shown in Figs. 5.1 and 5.2 for the SF and FF channels, respectively.

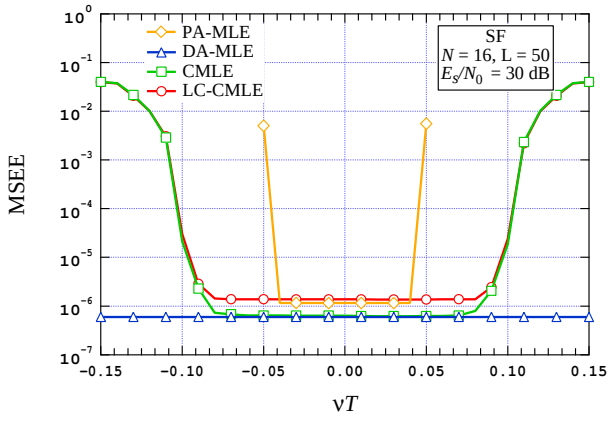

Figure 5.1: Acquisition range on slow-fading channel at $E_s/N_0 = 30$ dB.

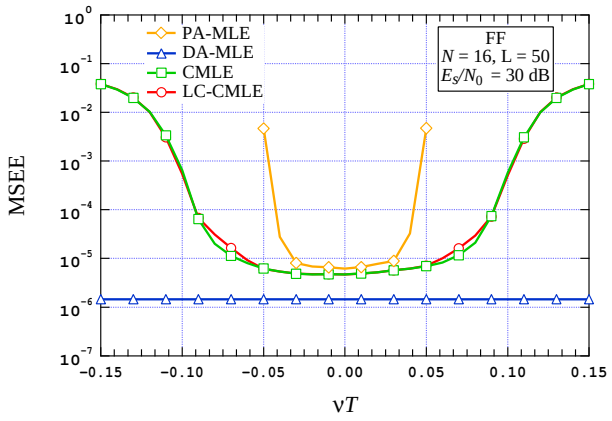

Figure 5.2: Acquisition range on fast-fading channel at $E_s/N_0 = 30$ dB.

For both scenarios, *i*) the MSEE of the DA-MLE (in the considered interval) comes out to be nearly independent of the true CFO, and *ii*) the LC-CMLE outperforms the PA-MLE, even though exhibiting a comparable complexity. This behaviour is consistent with the fact that the acquisition range is roughly as wide as the inverse of the time-domain pilot spacing. Indeed, when we arrange pilots adjacent together as in DA-MLE, the acquisition range is equal to the inverse of symbol interval $1/T$, while spacing them every 10 data symbols as in the PA-MLE, the acquisition range gets narrower to around 10 % of $1/T$. On the other side, the joint use of pilots and differential decisions, which are related to data symbols inserted between each pair of consecutive pilots, makes the acquisition range of the CMLE and LC-CMLE around twice as wider as compared with the PA-MLE, thus balancing the cost of pilot overhead against the acquisition performance. When the channel becomes faster, the acquisition performance of all the algorithms we are discussing slightly degrades as shown in Fig. 5.2. This takes place for both the CMLE and LC-CMLE as well. Actually, the larger the Doppler bandwidth, or in other words, the less the fading correlation over time and frequency, the less accurate the assumption a1) of Sect. 5.2.2 is. This leads to larger errors on the differential decisions from the products $z_l^{(s)} z_{l+k}^{(c)*}$ in (5.18), with the ultimate effect that the CFO estimate accuracy decays due to a less and less precise cancellation of the modulation phase shift in the metric (5.19).

Noise sensitivity. Figs. 5.3 and 5.4 quantify the MSEE metric attained over the SF and FF channels, respectively, as a function of E_s/N_0 assuming that the CFO to be estimated is $\nu = 0$.

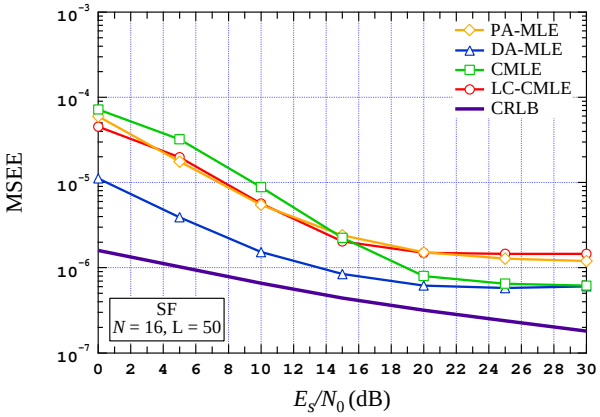

Figure 5.3: MSEE versus E_s/N_0 on slow-fading channel.

As for the SF scenario, from Fig. 5.3 it can be noted that: *i*) the DA-MLE does not achieve the CRLB due to the finite data records, whereas the additional MSEE gap of the

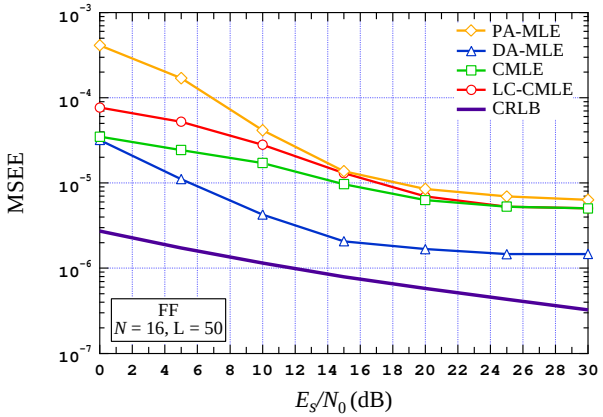

Figure 5.4: MSEE versus E_s/N_0 on fast-fading channel.

other estimators depends, respectively, on the limited number of pilots transmitted within the burst (PA-MLE), the errors on the differential decisions (CMLE) and the approximation on the fading covariance matrix (LC-CMLE) as a diagonal block matrix; *ii*) as expected, the CMLE closely approaches the DA-MLE at large SNRs thanks to all the information it exploits within the burst; *iii*) the LC-CMLE offers worse performance with respect to the CMLE because of the simplification of the fading covariance matrix structure, while it is substantially equivalent to the PA-MLE. Finally, the results for fast channel conditions (FF) are illustrated in Fig. 5.4. Due to the faster variations of fading from symbol to symbol, the MSEE curves of all the estimators depart slightly further from the CRLB if compared with the SF channel. Then, coherently with the results in Fig. 5.2, the reduced reliability of the differential decisions makes now the CMLE and the LC-CMLE incur a MSEE floor at large SNRs, comparable with that of the PA-MLE, although at low SNRs the CMLE proves to be the estimator closest to the DA-MLE.

Bibliography

[1] B. Crow, I. Widjaja, and I. Jeanclaude, "IEEE 802.11 wireless local area networks," *IEEE Commun. Mag.*, vol. 35, pp. 116–126, September 1997. 2

[2] C. Eklund, R. B. Marks, K. L. Stanwood, and S. Wang, "IEEE standard 802.16: A technical overview of the wirelessMAN air interface for broadband wireless access," *IEEE Commun. Mag.*, pp. 98–107, June 2002. 2, 20

[3] K. I. Pedersen, T. E. Kolding, F. Frederiksen, I. Z. Kovacs, D. Laselva, and P. E. Moegensen, "An overview of downlink radio resource management for UTRAN LTE," *IEEE Commun. Mag.*, vol. 47, no. 7, pp. 86–93, July 2009. 2, 20, 39

[4] J. A. C. Bingham, "Multicarrier modulation for data transmission: An idea whose time has come," *IEEE Commun. Mag.*, vol. 28, no. 5, pp. 5–14, May 1990. 2

[5] G. Cherubini, E. Eleftheriou, and S. Olcer, "Filtered multitone modulation for high-speed digital subscribers lines," *IEEE J. Select. Areas Comm.*, vol. 20, no. 5, pp. 1016–1028, June 2002. 3, 11

[6] M. Nouri, V. Lottici, R. Reggiannini, D. Ball, and M. Rayne, "TEDS: A high speed digital mobile communcation air interface for professional users," *IEEE Veh. Tecnhol. Mag.*, vol. 1, no. 4, pp. 32–42, December 2006. 3

[7] *ETSI EN 301 958, Digital Video Broadcasting (DVB): Interaction Channel for Digital Terrestrial Television (RCT) Incorporating Multiple Access OFDM*, Std., March 2002. 3

[8] *ETSI TS 300 392-2 V 3.1.1, Terrestrial Trunked Radio (TETRA): TETRA Enhanced Data Services (TEDS), Air Interface Specification*, Std., September 2006. 3

[9] J. G. Proakis, *Digital Communications*, 5th ed. McGraw Hill Higher Education, January 2008. 8

[10] P. A. Bello, "Characterization of randomly time-varying linear channels," *IEEE Trans. Commun. Syst.*, vol. 11, pp. 360–393, August 1963. 8

[11] R. A. Horn and C. R. Johnson, *Matrix Analysis*. New York: Cambridge University Press, 1985. 8

[12] P. Lancaster, *Theory of Matrices*. New York: Academic Press, 1969. 8

[13] V. Lottici, M. Luise, C. Saccomando, and F. Spalla, "Blind carrier frequency tracking for filterbank multicarrier wirelss communications," *IEEE Trans. Commun.*, vol. 53, no. 9, pp. 1762–1772, September 2005. 11, 60

[14] T. Keller and L. Hanzo, "Adaptive multicarrier modulation: A convenient framework for time-frequency processing in wireless communications," *IEEE Proc. of the IEEE*, vol. 88, no. 5, pp. 611–640, May 2000. 13

[15] P. S. Chow, J. M. Cioffi, and J. A. C. Bingham, "A practical discrete multitone transceiver loading algorithm for data transmission over spectrally shaped channels," *IEEE Trans. Commun.*, vol. 43, pp. 773–775, February/March/April 1995. 16

[16] R. Fischer and J. Huber, "A new loading algorithm for discrete multitone transmission," in *Proc. IEEE Globecom 1996*, vol. 1, 1996, pp. 724–728.

[17] S. Catreux, V. Erceg, D. Gesbert, and R. W. Heath, "Adaptive modulation and MIMO coding for broadband wireless data networks," *IEEE Commun. Mag.*, vol. 40, no. 6, pp. 108–115, June 2002. 13

[18] T. Cover and J. Thomas, *Elements of Information Theory*. Wiley, New York, 1991. 15

[19] G. J. Foschini and J. Salz, "Digital communications over fading radio channnels," *Bell Syst. Tech. J.*, pp. 429–456, February 1983. 15

[20] A. J. Goldsmith and S.-G. Chua, "Variable-rate variable-power MQAM for fading channels," *IEEE Trans. Commun.*, vol. 45, no. 10, pp. 1218–1230, October 1997. 15

[21] I. Kalet, "The multitone channel," *IEEE Trans. Commun.*, vol. 37, pp. 119–124, February 1989. 16

[22] J. Campello, "Practical bit loading for dmt," in *IEEE Proc. Int. Conf. Commun. 1999 (ICC '99)*, Vancouver, BC, Canada, June 1999, pp. 801–805. 16

[23] T. J. Willink and P. H. Wittke, "Optimization and performance evaluation of multicarrier transmission," *IEEE Trans. Inf. Theory*, vol. 43, pp. 426–440, March 1997. 16

[24] A. M. Wyglinski, F. Labeau, and P. Kabal, "Bit loading with BER-constraint for multicarrier systems," *IEEE Trans. Wireless Commun.*, vol. 4, pp. 1383–1387, July 2005. 16

[25] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004. 16, 22, 26, 29, 49

[26] C. Y. Wong, R. S. Cheng, K. B. Letaief, and R. D. Murch, "Multiuser OFDM with adaptive subcarrier, bit and power allocation," *IEEE J. Select. Areas Comm.*, vol. 17, no. 10, pp. 1747–1758, October 1999. 18, 29, 31, 32, 38

[27] I. C. Wong, Z. Shen, B. L. Evans, and J. G. Andrews, "A low complexity algorithm for proportional resource allocation in OFDMA systems," in *Proc. IEEE Workshop on Signal Processing Systems*, October 2004, pp. 1–6.

[28] D. Kivanc, G. Li, and H. Liu, "Computationally efficient bandwidth allocation for power control for OFDMA," *IEEE Trans. Wireless Commun.*, vol. 2, pp. 1150–1158, November 2003. 18

[29] H. Yin and H. Liu, "An efficient multiuser loading algorithm for OFDM-based broadband wireless systems," in *Proc. IEEE Globecom 2000*, vol. 1, November 2000, pp. 103–107. 18

[30] G. Song and Y. G. Li, "Cross-layer optimization for OFDM wireless networs-part I: Theoretical framwork," *IEEE Trans. Wireless Commun.*, vol. 4, pp. 614–624, March 2005.

[31] ——, "Cross-layer optimization for OFDM wireless networs-part II: Algorithm development," *IEEE Trans. Wireless Commun.*, vol. 4, pp. 625–634, March 2005. 18

[32] J. Jang and K. B. Lee, "Transmit power adaptation for multiuser OFDM systems," *IEEE J. Select. Areas Comm.*, vol. 21, pp. 171–178, February 2003. 18

[33] W. Rhee and J. M. Cioffi, "Increase in capacity of multiuser OFDM system using dynamic subchannel allocation," in *Proc. IEEE VTC 2000 Spring*, Tokio, Japan, May 2000, pp. 1085–1089. 23

[34] Z. Shen, J. G. Andrews, and B. L. Evans, "Optimal power allocation in multiuser OFDM systems," in *Proc. IEEE Globecom 2003*, vol. 1, December 2003, pp. 337–341. 18

[35] W. Yu and R. Lui, "Dual methods for nonconvex spectrum optimization of multicarrier systems," *IEEE Trans. Comm.*, vol. 54, no. 7, pp. pp. 1310–1322, July 2006. 21, 38, 49

[36] K. Seong, M. Mohseni, and J. M. Cioffi, "Optimal resource allocation for OFDMA downlink systems," in *Proc. IEEE Int Information Theory Symp*, 2006, pp. 1394–1398. 22

[37] Z. Shen, J. G. Andrews, and B. L. Evans, "Adaptive resource allocation in multiuser OFDM systems with proportional rate constraints," *IEEE Trans. Wireless Commun.*, vol. 4, pp. 2726–2737, November 2005. 25

[38] R. Baldick, *Applied Optimization: Formulation and Algorithms for Engineering Systems*. Cambridge University Press, 2006. 26

[39] T. . v8.4.0, "Evolved universal terrestrial radio access (E-UTRA): Physical channels and modulation," 3GPP, Tech. Rep., September 2008. 27

[40] J. Tang and X. Zhang, "Cross-layer design of dynamic resource allocation with diverse QoS guarantees for MIMO-OFDM wireless networks," in *Proc. Sixth IEEE Int. Symp. a World of Wireless Mobile and Multimedia Networks WoWMoM 2005*, 2005, pp. 205–212. 27

[41] I. Kim, I. Park, and Y. Lee, "Use of linear programming for dynamic subcarrier and bit allocation in multiuser OFDM," *IEEE Trans. Veh. Technol.*, vol. 55, no. 4, pp. 1195–1207, July 2006. 31

[42] L. G. Khachian, "A polynomial time algorithm in linear programming," *U.S.S.R. Comput. Math. and Math. Phys.*, vol. 20, pp. 53–72, 1980. 32, 42

[43] M. Moretti and A. Todini, "A reduced complexity cross-layer radio resource allocator for OFDMA systems," *IEEE Trans. Wireless Commun.*, vol. 6, pp. 2807–2812, 2007. 32, 39

[44] A. Abrardo, A. Alessio, P. Detti, and M. Moretti, "Centralized radio resource allocation for OFDMA cellular systems," in *Proc. IEEE ICC 2007*, 2007. 32

[45] R. K. Ahuja, T. L. Magnanti, and J. B. Orlin, *Network Flows*. Prentice Hall, 1993. 32, 42

[46] D. Bertsimas and J. Tsitsiklis, *Introduction to Linear Optimization*. Athena Scientific, 1998. 32, 42

[47] A. Todini, M. Moretti, A. Valletta, and A. Baiocchi, "A modular cross-layer scheduling and resource allocation architecture for OFDMA systems," in *Proc. IEEE Globecom 2006*, November 2006. 32, 36, 39, 47

[48] T. Ojanpera and R. Prasad, *Wideband CDMA for Third Generation Mobile Communications.* Artech House, 2001. 35

[49] LG, "Standard aspects of interference coordination for EUTRA," 3GPP R1-051051, Tech. Rep., 2005. 35

[50] Z.-Q. Luo and S. Zhang, "Dynamic spectrum management: Complexity and duality," *IEEE J. Select. Topics Signal Process.,*, vol. 2, no. 1, pp. 57–73, February 2008. 35

[51] L. Venturino, N. Prasad, and X. Wang, "Coordinated scheduling and power allocation in downlink multicell OFDMA networks," *IEEE Trans. Veh. Technol.*, vol. 58, no. 6, pp. 2835–2848, July 2009. 35

[52] G. Li and H. Liu, "Downlink radio resource allocation for multi-cell OFDMA systems," *IEEE Trans. Wireless Commun.*, vol. 5, no. 12, pp. 3451–3459, December 2006. 35

[53] S. Gault, W. Hachem, and P. Ciblat, "Performance analysis of an OFDMA transmission system in a multicell environment," *IEEE Trans. Commun.*, vol. 55, no. 4, pp. 740–751, April 2007. 35

[54] A. Abrardo, A. Alessio, P. Detti, and M. Moretti, "Radio resource allocation problems for OFDMA cellular systems," *Comput. Oper. Res.*, vol. 36, no. 5, pp. 1572–1581, 2009. 36, 52

[55] M. Pischella and J. C. Belfiore, "Distributed resource allocation for rate-constrained users in multi-cell OFDMA networks," *IEEE Commun. Lett.*, vol. 12, no. 4, pp. 250–252, April 2008. 36, 52

[56] R. Y. Chang, Z. Tao, J. Zhang, and C. C. J. Kuo, "A graph approach to dynamic fractional frequency reuse (FFR) in multi-cell OFDMA networks," in *Proc. IEEE Int. Conf. Commun. 2009 (ICC '09)*, 2009, pp. 1–6. 36, 53

[57] W. Yu, W. Rhee, S. Boyd, and J. Cioffi, "Iterative water-filling for gaussian vector multiple-access channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 1, pp. 145–152, Jan. 2004. 38

[58] B. Korte and J. Vygen, *Combinatorial Optimiza.* Springer, 2008. 41

[59] *TS 36.420, Evolved Universal Terrestrial Radio Access Network (E-UTRAN): X2 general aspects and principles*, 3GPP-LTE Std., 2009. 42

[60] M. Rahman and H. Yanikomeroglu, "Enhancing cell-edge performance: a downlink dynamic interference avoidance scheme with inter-cell coordination," *IEEE Trans. Wireless Commun.*, vol. 9, no. 4, pp. 1414–1425, 2010. 42

[61] R. Yates, "A framework for uplink power control in cellular radio systems," *IEEE J. Select. Areas Comm.*, vol. 13, no. 7, pp. 1341–1347, Sep 1995. 44

[62] B. Bensaou, D. Tsang, and K. T. Chan, "Credit-based fair queueing (CBFQ): a simple service-scheduling algorithm for packet-switched networks," *IEEE/ACM Trans. Netw.*, vol. 9, no. 5, pp. 591–604, Oct. 2001. 46

[63] A. Todini and A. Baiocchi, "Scheduling algorithm," Infocom Department, University of Rome "La Sapienza", Tech. Rep., June 2010. 46

[64] J. Papandriopoulos and J. S. Evans, "SCALE: A low-complexity distributed protocol for spectrum balancing in multiuser DSL networks," *IEEE Trans. Inf. Theory*, vol. 55, no. 8, pp. 3711–3724, Aug. 2009. 48

[65] J. Zander, "Performance of optimum transmitter power control in cellular radio systems," *IEEE Trans. Veh. Technol.*, vol. 41, no. 1, pp. 57–62, Feb. 1992. 50

[66] T. Ihalainen, T. H. Stitz, M. Rinne, and M. Renfors, "Channel equalization in filter bank based multicarrier modulation for wireless communications," *EURASIP Journal on Advances in Signal Process.*, vol. 2007, p. 18, 2007. 59

[67] L. Vandendorpe, L. Cuvelier, F. Deryck, J. Louveaux, and O. van de Wiel, "Fractionally spaced linear and decision feedback detectors for transmultiplexers," *IEEE Trans. Signal Process.*, vol. 46, no. 4, pp. 996–1011, April 1998. 59

[68] J. Louveaux, L. Vandendorpe, L. Cuvelier, and T. Pollet, "Bit-rate sensitivity of filterbank-based VDSL transmission to timing errors," *IEEE Trans. Commun.*, vol. 49, no. 2, pp. 375–384, February 2001. 59

[69] ——, "An early-late timing recovery scheme for filter-bank-based multicarrier transmission," *IEEE Trans. Commun.*, vol. 48, no. 48, pp. 1746–1754, October 2000. 59

[70] J. Louveaux, L. Cuvelier, L. Vandendorpe, and T. Pollet, "Baud rate timing recovery scheme for filter-bank-based multicarrier transmission," *IEEE Trans. Commun.*, vol. 51, no. 4, pp. 652–663, April 2003. 59

[71] V. Lottici, M. Luise, C. Saccomando, and F. Spalla, "Non-data-aided timing recovery for filter-bank multicarrier wireless communications," *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4365–4375, November 2006. 59

[72] M. G. Hebley and D. P. Taylor, "The effect of diversity on a burst-mode carrier frequency estimator in the frequency-selective multipath channel," *IEEE Trans. Commun.*, vol. 46, no. 4, pp. 553–560, April 1998. 59

[73] M. Morelli and U. Mengali, "Carrier-frequency estimation for transmission over selective channels," *IEEE Trans. Commun.*, vol. 48, no. 9, pp. 1580–1589, September 2000.

[74] P. H. Moose, "A technique for OFDM frequency offset correction," *IEEE Trans. Commun.*, vol. 42, no. 10, pp. 2908–2914, 1994.

[75] J. J. V. de Beek, M. Sandell, and P. O. Borjesson, "ML estimation of time and frequency offset in OFDM systems," *IEEE Trans. Signal Process.*, vol. 45, no. 7, pp. 1800–1805, July 1997.

[76] T. Lv, H. Li, and J. Chen, "Joint estimation of symbol timing and carrier frequency offset of OFDM signals over fast time-varying multipath channels," *IEEE Trans. Signal Process.*, vol. 53, no. 12, pp. 4526–4535, December 2005.

[77] Y. Jiang, H. Minn, X. Gao, X. You, and Y. Li, "Frequency offset estimation and training sequence design for MIMO OFDM," *IEEE Trans. Wireless Commun.*, vol. 7, no. 4, pp. 1244–1254, April 2008. 59

[78] V. Lottici, R. Reggiannini, and M. Carta, "Pilot-aided carrier frequency estimation for filter-bank multicarrier wireless communications on doubly-selective channels," *IEEE Trans. Signal Process.*, vol. 58, no. 5, pp. 2783–2794, May 2010. 60, 63

[79] T. Fusco, A. Petrella, and M. Tanda, "Blind CFO estimation for noncritically sampled FMT systems," *IEEE Trans. Signal Process.*, vol. 56, no. 6, pp. 2603–2608, June 2008. 60

[80] H. Bölcskei, "Blind estimation of symbol and carrier frequency offset in wireless OFDM systems," *IEEE Trans. Commun.*, vol. 49, no. 6, pp. 988–999, June 2001. 60

[81] T. Fusco, A. Petrella, and M. Tanda, "Data-aided symbol timing and CFO synchronization for filter bank multicarrier systems," *IEEE Trans. Wireless Commun.*, vol. 8, no. 5, pp. 2705–2715, May 2009. 60

[82] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory.* Englewood Cliffs, NJ: Prentice-Hall, 1993. 63, 66

Biography

Giulio Dainelli was born in Empoli, Italy, on June 19, 1983.

In 2002 he received the *diploma di maturità scientifica* from Liceo Scientifico Statale "Il Pontormo" in Empoli, Italy.

Since 2002 he joined the Faculty of Engineering of the University of Pisa, from which he received the Doctor Engineer degree (cum laude) in Telecommunication Engineering in 2007.

Since 2008 he was with the Department of Information Engineering of the University of Pisa, where he worked toward the Ph.D. degree in Information Engineering under the supervision of Prof. Aldo Nunzio D'Andrea, Ing. Marco Moretti and Prof. Michele Morelli.

Research Interests

His expertise and general interests span the areas of wireless communications and signal processing, estimation theory. Current research topics focus on resource allocation for multicarrier cellular OFDMA systems, MIMO systems, synchronization for FBMC communications.